TRB

TRANSPORTATION RESEARCH RECORD

Journal of the Transportation Research Board, No. 2159

Railways
2010

TRANSPORTATION RESEARCH BOARD
OF THE NATIONAL ACADEMIES

TRANSPORTATION RESEARCH BOARD
OF THE NATIONAL ACADEMIES

TRANSPORTATION RESEARCH RECORDS, which are published throughout the year, consist of collections of papers on specific transportation modes and subject areas. Each Record is classified according to the subscriber category or categories covered by the papers published in that volume. The views expressed in papers published in the Transportation Research Record series are those of the authors and do not necessarily reflect the views of the peer review committee(s), the Transportation Research Board, the National Research Council, or the sponsors of TRB activities. The Transportation Research Board does not endorse products or manufacturers; trade and manufacturers' names may appear in a Record paper only if they are considered essential.

PEER REVIEW OF PAPERS: All papers published in the Transportation Research Record series have been reviewed and accepted for publication through the Transportation Research Board's peer review process established according to procedures approved by the Governing Board of the National Research Council. Papers are refereed by the TRB standing committees identified on page ii of each Record. Reviewers are selected from among committee members and other outside experts. The Transportation Research Board requires a minimum of three reviews; a decision is based on reviewer comments and resultant author revision. For details about the peer review process, see the information on the inside back cover.

THE TRANSPORTATION RESEARCH RECORD PUBLICATION BOARD, comprising a cross section of transportation disciplines and with equal representation from the academic and practitioner communities, assures that the quality of the *Transportation Research Record: Journal of the Transportation Research Board* and the TRB paper peer review process is consistent with that of a peer-reviewed scientific journal. Members from the academic community make decisions on granting tenure; all academic and practitioner members have authored papers published in peer-reviewed journals and all have participated in the TRB peer review process.

TRANSPORTATION RESEARCH BOARD PUBLICATIONS may be ordered directly from the TRB Business Office, through the Internet at www.TRB.org, or by annual subscription through organizational or individual affiliation with TRB. Affiliates and library subscribers are eligible for substantial discounts. For further information, contact the Transportation Research Board Business Office, 500 Fifth Street, NW, Washington, DC 20001 (telephone 202-334-3213; fax 202-334-2519; or e-mail TRBsales@nas.edu).

TRANSPORTATION RESEARCH RECORD PAPERS ONLINE: The TRR Journal Online website provides electronic access to the full text of papers that have been published in the Transportation Research Record series since 1996. The site is updated as new Record papers become available. To search abstracts and to find subscription and pricing information, go to www.TRB.org/TRROnline.

TRANSPORTATION RESEARCH BOARD 2010 EXECUTIVE COMMITTEE*

Chair: Michael R. Morris, Director of Transportation, North Central Texas Council of Governments, Arlington
Vice Chair: Neil J. Pedersen, Administrator, Maryland State Highway Administration, Baltimore
Executive Director: Robert E. Skinner, Jr., Transportation Research Board

J. Barry Barker, Executive Director, Transit Authority of River City, Louisville, Kentucky
Allen D. Biehler, Secretary, Pennsylvania Department of Transportation, Harrisburg
Larry L. Brown, Sr., Executive Director, Mississippi Department of Transportation, Jackson
Deborah H. Butler, Executive Vice President, Planning, and CIO, Norfolk Southern Corporation, Norfolk, Virginia
William A. V. Clark, Professor, Department of Geography, University of California, Los Angeles
Eugene A. Conti, Jr., Secretary of Transportation, North Carolina Department of Transportation, Raleigh
Nicholas J. Garber, Henry L. Kinnier Professor, Department of Civil Engineering, and Director, Center for Transportation Studies, University of Virginia, Charlottesville
Jeffrey W. Hamiel, Executive Director, Metropolitan Airports Commission, Minneapolis, Minnesota
Paula J. Hammond, Secretary, Washington State Department of Transportation, Olympia
Edward A. (Ned) Helme, President, Center for Clean Air Policy, Washington, D.C.
Adib K. Kanafani, Cahill Professor of Civil Engineering, University of California, Berkeley (Past Chair, 2009)
Susan Martinovich, Director, Nevada Department of Transportation, Carson City
Debra L. Miller, Secretary, Kansas Department of Transportation, Topeka (Past Chair, 2008)
Sandra Rosenbloom, Professor of Planning, University of Arizona, Tucson
Tracy L. Rosser, Vice President, Corporate Traffic, Wal-Mart Stores, Inc., Mandeville, Louisiana
Steven T. Scalzo, Chief Operating Officer, Marine Resources Group, Seattle, Washington
Henry G. (Gerry) Schwartz, Jr., Chairman (retired), Jacobs/Sverdrup Civil, Inc., St. Louis, Missouri
Beverly A. Scott, General Manager and Chief Executive Officer, Metropolitan Atlanta Rapid Transit Authority, Atlanta, Georgia
David Seltzer, Principal, Mercator Advisors LLC, Philadelphia, Pennsylvania
Daniel Sperling, Professor of Civil Engineering and Environmental Science and Policy; Director, Institute of Transportation Studies; and Interim Director, Energy Efficiency Center, University of California, Davis
Kirk T. Steudle, Director, Michigan Department of Transportation, Lansing
Douglas W. Stotlar, President and Chief Executive Officer, Con-Way, Inc., Ann Arbor, Michigan
C. Michael Walton, Ernest H. Cockrell Centennial Chair in Engineering, University of Texas, Austin (Past Chair, 1991)

Peter H. Appel, Administrator, Research and Innovative Technology Administration, U.S. Department of Transportation (ex officio)
J. Randolph Babbitt, Administrator, Federal Aviation Administration, U.S. Department of Transportation (ex officio)
Rebecca M. Brewster, President and COO, American Transportation Research Institute, Smyrna, Georgia (ex officio)
George Bugliarello, President Emeritus and University Professor, Polytechnic Institute of New York University, Brooklyn; Foreign Secretary, National Academy of Engineering, Washington, D.C. (ex officio)
Anne S. Ferro, Administrator, Federal Motor Carrier Safety Administration, U.S. Department of Transportation (ex officio)
LeRoy Gishi, Chief, Division of Transportation, Bureau of Indian Affairs, U.S. Department of the Interior, Washington, D.C. (ex officio)
Edward R. Hamberger, President and CEO, Association of American Railroads, Washington, D.C. (ex officio)
John C. Horsley, Executive Director, American Association of State Highway and Transportation Officials, Washington, D.C. (ex officio)
David T. Matsuda, Deputy Administrator, Maritime Administration, U.S. Department of Transportation (ex officio)
Victor M. Mendez, Administrator, Federal Highway Administration, U.S. Department of Transportation (ex officio)
William W. Millar, President, American Public Transportation Association, Washington, D.C. (ex officio) (Past Chair, 1992)
Robert J. Papp (Adm., U.S. Coast Guard), Commandant, U.S. Coast Guard, U.S. Department of Homeland Security (ex officio)
Cynthia L. Quarterman, Administrator, Pipeline and Hazardous Materials Safety Administration, U.S. Department of Transportation (ex officio)
Peter M. Rogoff, Administrator, Federal Transit Administration, U.S. Department of Transportation (ex officio)
David L. Strickland, Administrator, National Highway Traffic Safety Administration, U.S. Department of Transportation (ex officio)
Joseph C. Szabo, Administrator, Federal Railroad Administration, U.S. Department of Transportation (ex officio)
Polly Trottenberg, Assistant Secretary for Transportation Policy, U.S. Department of Transportation (ex officio)
Robert L. Van Antwerp (Lt. General, U.S. Army), Chief of Engineers and Commanding General, U.S. Army Corps of Engineers, Washington, D.C. (ex officio)

* Membership as of August 2010.

TRB

TRANSPORTATION RESEARCH RECORD

Journal of the Transportation Research Board, No. 2159

Railways
2010

A Peer-Reviewed Publication

TRANSPORTATION RESEARCH BOARD
OF THE NATIONAL ACADEMIES

Washington, D.C.
2010

www.TRB.org

Transportation Research Record 2159
ISSN 0361-1981
ISBN 978-0-309-14286-1

Subscriber Categories
Railroads; passenger transportation; freight transportation; planning and forecasting

Printed in the United States of America

TRANSPORTATION RESEARCH RECORD PUBLICATION BOARD
C. Michael Walton, PhD, PE, Ernest H. Cockrell Centennial Chair in Engineering and Professor of Civil Engineering, University of Texas, Austin (Cochair)
Mary Lynn Tischer, PhD, Director, Office of Transportation Policy Studies, Federal Highway Administration, Washington, D.C. (Cochair)
Daniel Brand, SB, SM, PE, Consultant, Lyme, New Hampshire
Mary R. Brooks, BOT, MBA, PhD, William A. Black Chair of Commerce, Dalhousie University, Halifax, Nova Scotia, Canada
Charles E. Howard, Jr., MCRP, Director, Transportation Planning, Puget Sound Regional Council, Seattle, Washington
Thomas J. Kazmierowski, PEng, Manager, Materials Engineering and Research Office, Ontario Ministry of Transportation, Toronto, Canada
Michael D. Meyer, PhD, PE, Frederick R. Dickerson Professor, School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta
Sandra Rosenbloom, MA, PhD, Professor of Planning, University of Arizona, Tucson
Kumares C. Sinha, PhD, PE, Olson Distinguished Professor, School of Civil Engineering, Purdue University, West Lafayette, Indiana
L. David Suits, MSCE, Executive Director, North American Geosynthetics Society, Albany, New York

Peer Review of Transportation Research Record 2159

RAIL GROUP
Robert M. Dorer, Research and Innovative Technology Administration (Chair)

Intercity Passenger Rail Committee
Anthony D. Perl, Simon Fraser University (Chair), Eric S. Tyrer II, York College, City University of New York (Secretary), Rohit T. Aggarwala, Daniel Brand, William D. Bronte, Ross B. Capon, John F. Cikota, Jr., Allison Lee C. de Cerreno, Dharm Guruswamy, George Haikalis, Jason J. Maga, Ronald A. Mauri, Matthew James Melzer, Neil E. Moyer, Andrew B. Nash, Howard R. Permut, Charles H. Quandel, Sr., Daniel L. Roth, Joseph P. Schwieterman, David P. Simpson, Alan C. Tobias, John C. Tone, Camille Tsao, Vukan R. Vuchic, Randall E. Wade, Mark C. Walbrun, Shirley R. Williams, Frank N. Wilner, Albert C. Witzig

Passenger Rail Equipment and Systems Integration Committee
Patrick B. Simmons, North Carolina Department of Transportation (Chair), Eloy Martinez, Federal Railroad Administration (Secretary), John G. Bell, Alan J. Bing, Nora C. Friend, Andrew J. Galloway, John A. Harrison, Joseph Lee Hutchins, Jr., Larry David Kelterborn, Andrés López-Pita, Nobuyuki Matsumoto, David O. Nelson, Philip Olekszyk, R. Scott Phelan, Charles A. Poltenson, Sr., Sunduck Daniel Suh, Michael J. Trosino, David A. Valenstein, Robert B. Watson

Railroad Operating Technologies Committee
Gordon B. Mott, Jacksonville, Florida (Chair), Yung-Cheng Lai, National Taiwan University (Vice Chair), Adrian D. Hellman, Volpe National Transportation Systems Center (Secretary), Virginia M. Beck, Stephen Juan Bruno, Olga K. Cataldi, Richard U. Cogswell, Llewellyn C. Davis, Timothy J. DePaepe, Mark H. Dingler, Jeremiah Dirnberger, Mark W. Hemphill, Ecton K. Holt, Michael Iden, Edwin R. Kraft, Coleman Lawrence, Robert H. Leilich, Marco M. Luethi, Jena C. Montgomery, Joern Pachl, Paul H. Reistrup, Mark K. Ricci, Ismail Sahin, Victor Simuoli, Michael E. Smith, Mark P. Stehly, James A. Stem, Jr.

Local and Regional Rail Freight Transport Committee
David B. Clarke, University of Tennessee, Knoxville (Chair), Gary E. Landrio, TranSystems Corporation (Secretary), Tim Andrews, C. Howard Capito, Diane Davidson, John M. Gibson, Jr., John T. Gray II, George Avery Grimes, Donald B. Ludlow, Richard A. Nordahl, R. Leo Penne, Charles A. Poltenson, Sr., Henry Posner III, George Raymond, Richard Durward Robey, James H. M. Savage, George W. Schafer III, David P. Simpson, Anne Strauss-Wieder, Peter F. Swan, Forrest R. Van Schwartz

Railroad Track Structure System Design Committee
David D. Davis, Transportation Technology Center, Inc. (Chair), Dwight W. Clark, Union Pacific Railroad Company (Secretary), Ernest J. Barenberg, Timothy R. Bennett, David N. Bilow, Randy L. Bowman, Michael O. Brown, Miodrag Budisa, William G. Byers, Laurence E. Daniels, Carlton L. Ho, Henry M. Lees, Jr., Andrés López-Pita, Dale W. Ophardt, Donald Plotkin, David E. Staplin, Yu-Jiang Zhang

Railway Maintenance Committee
Conrad J. Ruppert, Jr., National Railroad Passenger Corporation (Chair), James P. Hyslip, HyGround Engineering, LLC (Secretary), Miodrag Budisa, Gary A. Carr, Thomas F. DeJoseph, Willem Ebersohn, Riley Edwards, Gurmel S. Ghataora, P. J. Hannes Gräbe, Carlton L. Ho, Hai Huang, Stanislav Jovanovic, David L. Kannenberg, Francesco Lanza Di Scalea, Eric E. Magel, Valerie Marcolongo, Daniel B. Mesnick, Thomas H. O'Brien, Bruce Pohlot, Christian Mark Roberts, Michael D. Roney, Mario A. Ruel, Ernest T. Selig, Eric Sherrock, Theodore R. Sussmann, Jr., Vincent R. Terrill, Jackie van der Westhuizen, Timothy R. Wells, Allan M. Zarembski

FREIGHT SYSTEMS GROUP
Paul H. Bingham, IHS Global Insight, Inc. (Chair)

Freight Transportation Economics and Regulation Committee
Kevin Neels, Brattle Group (Chair), Catherine L. Taylor, Volpe National Transportation Systems Center (Secretary), Michael H. Belzer, Michael A. Benouaich, Stephen V. Burks, David B. Clarke, Deborah M. Freund, Konstantina Gkritza, John T. Gray II, Mark A. Hornung, John H. Siebert, Michael E. Smith, Joseph A. Swanson, Louis S. Thompson, Steven J. Tripp, Karen E. White, Peter Wolff

Transportation of Hazardous Materials Committee
Mark Lepofsky, Visual Risk Technologies, Inc. (Chair), Richard C. Bornhorst, U.S. Coast Guard (Secretary), Mark D. Abkowitz, Laura L. Barhydt, Paul M. Bomgardner, Christopher J. Bonanti, Cheryl A. Burke, Timothy Butters, Edward R. Chapman, John L. Conley, John V. Currie, Sarah J. Dammen, Philip Joseph Daum, Ronald J. Duych, Robert E. Fronczak, Emily K. Goodenough, Scott C. Gorton, Charles H. Hochman, Joseph Gary Lanthrum, Richard Moskowitz, Steven L. Niswander, Mohd Rapik Saat, John Salvesen, William S. Schoonover, James O'Neal Simmons, Stephen F. Urschel, Robert D. Waters

Peer review is indicated by a footnote at the end of each paper. The organizational units, officers, and members are as of December 31, 2009.

Transportation Research Board Staff
Elaine King, Rail Transport Specialist
Matthew A. Miller, Senior Program Associate

Publications Office
Michael K. Hayes, Editor; Carol Jaka, Production Editor; Jennifer Correro, Proofreader; Claudia Sauls, Manuscript Preparer

Ann E. Petty, Managing Editor; Juanita Green, Production Manager; Phyllis Barber, Publishing Administrator; Jennifer J. Weeks, Manuscript Preparation Manager

THE NATIONAL ACADEMIES

Advisers to the Nation on Science, Engineering, and Medicine

The **National Academy of Sciences** is a private, nonprofit, self-perpetuating society of distinguished scholars engaged in scientific and engineering research, dedicated to the furtherance of science and technology and to their use for the general welfare. On the authority of the charter granted to it by the Congress in 1863, the Academy has a mandate that requires it to advise the federal government on scientific and technical matters. Dr. Ralph J. Cicerone is president of the National Academy of Sciences.

The **National Academy of Engineering** was established in 1964, under the charter of the National Academy of Sciences, as a parallel organization of outstanding engineers. It is autonomous in its administration and in the selection of its members, sharing with the National Academy of Sciences the responsibility for advising the federal government. The National Academy of Engineering also sponsors engineering programs aimed at meeting national needs, encourages education and research, and recognizes the superior achievements of engineers. Dr. Charles M. Vest is president of the National Academy of Engineering.

The **Institute of Medicine** was established in 1970 by the National Academy of Sciences to secure the services of eminent members of appropriate professions in the examination of policy matters pertaining to the health of the public. The Institute acts under the responsibility given to the National Academy of Sciences by its congressional charter to be an adviser to the federal government and, on its own initiative, to identify issues of medical care, research, and education. Dr. Harvey V. Fineberg is president of the Institute of Medicine.

The **National Research Council** was organized by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and advising the federal government. Functioning in accordance with general policies determined by the Academy, the Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in providing services to the government, the public, and the scientific and engineering communities. The Council is administered jointly by both the Academies and the Institute of Medicine. Dr. Ralph J. Cicerone and Dr. Charles M. Vest are chair and vice chair, respectively, of the National Research Council.

The **Transportation Research Board** is one of six major divisions of the National Research Council. The mission of the Transportation Research Board is to provide leadership in transportation innovation and progress through research and information exchange, conducted within a setting that is objective, interdisciplinary, and multimodal. The Board's varied activities annually engage about 7,000 engineers, scientists, and other transportation researchers and practitioners from the public and private sectors and academia, all of whom contribute their expertise in the public interest. The program is supported by state transportation departments, federal agencies including the component administrations of the U.S. Department of Transportation, and other organizations and individuals interested in the development of transportation. **www.TRB.org**

www.national-academies.org

TRB SPONSORS*

Transportation Departments of the 50 States, Puerto Rico, and the District of Columbia

Federal Government
U.S. Department of Transportation
- Federal Aviation Administration
- Federal Highway Administration
- Federal Motor Carrier Safety Administration
- Federal Railroad Administration
- Federal Transit Administration
- National Highway Traffic Safety Administration
- Research and Innovative Technology Administration

Bureau of Indian Affairs
Science and Technology Directorate, U.S. Department of Homeland Security
U.S. Army Corps of Engineers
U.S. Coast Guard

Nongovernmental Organizations
American Association of State Highway and Transportation Officials
American Public Transportation Association
American Transportation Research Institute
Association of American Railroads

---

*As of August 2010.

TRANSPORTATION RESEARCH RECORD

Journal of the Transportation Research Board, No. 2159

Contents

Foreword	**vii**
Application of Postassessment of Kyushy Shinkansen Network to Proposed U.S. High-Speed Railway Project Yuki Tanaka and Masaku Monji	1
High-Speed Rail Project Development Processes in the United States and China Xueming (Jimmy) Chen and Ming Zhang	9
Privatization Versus Public Works for High-Speed Rail Projects Tsung-Chung Kao, Yung-Cheng (Rex) Lai, and Mei-Cheng Shih	18
Energy Consumption and Emissions of High-Speed Trains Alberto García Álvarez	27
Genetic Algorithm–Based Column Generation Approach to Passenger Rail Crew Scheduling Mindy Liu, Ali Haghani, and Shahabeddin Toobaie	36
Deploying Business Logic for Railway Infrastructure Charging Rosário Macário, Paulo Teixeira, and Werner Rothengatter	44
New Approach to Appraisal of Rail Freight Projects in South Korea: Using the Value of Freight Transit Time Savings Kyungwoo Kang, Anne Strauss-Wieder, and Jin Ki Eom	52
Effect of Train Speed on Risk Analysis of Transporting Hazardous Materials by Rail Athaphon Kawprasert and Christopher P. L. Barkan	59

Optimizing Train Network Routing with Heterogeneous Traffic 69
Yung-Cheng (Rex) Lai, Mark H. Dingler, Chung-En Hsu,
and Pei-Chun Chiang

Effects of Communications-Based Train Control and Electronically Controlled Pneumatic Brakes on Railroad Capacity 77
Mark H. Dingler, Yung-Cheng (Rex) Lai, and Christopher P. L. Barkan

---

Force Characteristics of Longitudinally Coupled Slab Track Turnout on Bridges Under Temperature Action 85
JuanJuan Ren, Rui Xiang, and Xueyi Liu

Stress Dependence of Ultrasonic Guided Waves in Rails 91
Ivan Bartoli, Robert Phillips, Stefano Coccia, Ankit Srivastava,
Francesco Lanza di Scalea, Mahmood Fateh, and Gary Carr

Nondestructive Identification of Freezing-Induced Cracks in Concrete Sleepers of High-Speed Railways in South Korea 98
Sung-Ho Joh, Seon Keun Hwang, Tae-Ho Kang,
Chul-Soo Park, and Il-Wha Lee

Railroad Ballast Evaluation Using Ground-Penetrating Radar: Laboratory Investigation and Field Validation 110
Zhen Leng and Imad L. Al-Qadi

Foreword

The 2010 series of the *Transportation Research Record: Journal of the Transportation Research Board* consists of approximately 900 papers selected from 3,700 submissions after rigorous peer review. The peer review for each paper published in this volume was coordinated by the committee acknowledged at the end of the text; members of the reviewing committees for the papers in this volume are listed on page ii.

Additional information about the *Transportation Research Record: Journal of the Transportation Research Board* series and the peer review process appears on the inside back cover. TRB appreciates the interest shown by authors in offering their papers, and the Board looks forward to future submissions.

Note: Many of the photographs, figures, and tables in this volume have been converted from color to grayscale for printing. The electronic files of the papers, posted on the web at **www.TRB.org/TRROnline**, retain the color versions of photographs, figures, and tables as originally submitted for publication.

Measurement Conversion Factors

To convert from the unit in the first column to the unit in the second column, multiply by the factor in the third column.

Customary Unit	*SI Unit*	*Factor*
Length		
inches	millimeters	25.4
inches	centimeters	2.54
feet	meters	0.305
yards	meters	0.914
miles	kilometers	1.61
Area		
square inches	square millimeters	645.1
square feet	square meters	0.093
square yards	square meters	0.836
acres	hectares	0.405
square miles	square kilometers	2.59
Volume		
gallons	liters	3.785
cubic feet	cubic meters	0.028
cubic yards	cubic meters	0.765
Mass		
ounces	grams	28.35
pounds	kilograms	0.454
short tons	megagrams	0.907
Illumination		
footcandles	lux	10.76
footlamberts	candelas per square meter	3.426
Force and Pressure or Stress		
poundforce	newtons	4.45
poundforce per square inch	kilopascals	6.89
Temperature		
To convert Fahrenheit temperature (°F) to Celsius temperature (°C), use the following formula: °C = (°F – 32)/1.8		

SI Unit	*Customary Unit*	*Factor*
Length		
millimeters	inches	0.039
centimeters	inches	0.394
meters	feet	3.281
meters	yards	1.094
kilometers	miles	0.621
Area		
square millimeters	square inches	0.00155
square meters	square feet	10.764
square meters	square yards	1.196
hectares	acres	2.471
square kilometers	square miles	0.386
Volume		
liters	gallons	0.264
cubic meters	cubic feet	35.314
cubic meters	cubic yards	1.308
Mass		
grams	ounces	0.035
kilograms	pounds	2.205
megagrams	short tons	1.102
Illumination		
lux	footcandles	0.093
candelas per square meter	footlamberts	0.292
Force and Pressure or Stress		
newtons	poundforce	0.225
kilopascals	poundforce per square inch	0.145
Temperature		
To convert Celsius temperature (°C) to Fahrenheit temperature (°F), use the following formula: °F = (°C × 1.8) + 32		

Abbreviations Used Without Definitions

AASHO	American Association of State Highway Officials
AASHTO	American Association of State Highway and Transportation Officials
ACRP	Airport Cooperative Research Program
APTA	American Public Transportation Association
ASCE	American Society of Civil Engineers
ASTM	American Society for Testing and Materials (known by abbreviation only)
FAA	Federal Aviation Administration
FHWA	Federal Highway Administration
FMCSA	Federal Motor Carrier Safety Administration
FRA	Federal Railroad Administration
FTA	Federal Transit Administration
IEEE	Institute of Electrical and Electronics Engineers
ISO	International Organization for Standardization
ITE	Institute of Transportation Engineers
NASA	National Aeronautics and Space Administration
NCHRP	National Cooperative Highway Research Program
NHTSA	National Highway Traffic Safety Administration
RITA	Research and Innovative Technology Administration
SAE	Society of Automotive Engineers
SHRP	Strategic Highway Research Program
TCRP	Transit Cooperative Research Program
TRB	Transportation Research Board

Application of Postassessment of Kyushy Shinkansen Network to Proposed U.S. High-Speed Railway Project

Yuki Tanaka and Masaku Monji

In April 2009 the Obama administration announced its strategic plan for high-speed rail in the United States. That plan stipulates the public benefits that are to be expected from a high-speed railway system. This paper describes a recently published report outlining the postassessment of the Kyushu Shinkansen project conducted by the Japan Railway Construction, Transport, and Technology Agency. The paper focuses on the effects of the implementation of the Kyushu Shinkansen project on passenger movements and changes in business and environmental aspects that are relevant to the U.S. strategic plan. This is the first analysis outside Japan of the postassessment of the Kyushu Shinkansen project. It first examines the postassessment analysis of changes in traffic volume. It then addresses the effects of the Kyushu Shinkansen project as they relate to the public benefits expected from the U.S. high-speed rail strategic plan, including the economic and social effects as well as the environmental benefits. The paper attempts to provide the Japanese case as a reference for the U.S. high-speed railway project. The postassessment of the Japanese case indicates that the new Kyushu Shinkansen route has an influence on travel behavior when people use rail for commuting and tourism. It also addresses the critical factor of how aviation demand can be shifted to rail to gain substantial benefits. Taking these points into consideration, the Obama administration's *Vision for High-Speed Rail in America* is just a starting point. A wider network of railways with more vigorous funding must be expected.

On April 16, 2009, U.S. President Barack Obama called for high-speed passenger trains to transform travel in America and released a strategic plan outlining his vision for a high-speed rail (HSR) network in the United States. President Obama's *Vision for High-Speed Rail in America* (referred to here as the vision) (*1*), a strategic plan for the implementation of HSR, helps address the nation's transportation challenges by investing in an efficient, high-speed passenger rail network of intercity corridors across the United States. The federal funds allocated amount to $8 billion from the American Recovery and Reinvestment Act of 2009 and a $5 billion request in the president's budget. The administration also urges states and local communities to put together plans for a network of 100- to 600-mi corridors.

Y. Tanaka, Japan International Transport Institute, Institution for Transport Policy Studies, 3-18-19, Toranomon, Minato-ku, Tokyo 105-0001. M. Monji, Japan Railway Construction, Transport, and Technology Agency, 6-50-1, Honchou, Naka-ku, Yokohama City, Kanagawa 231-8315, Japan. Corresponding author: Y. Tanaka, y-tanaka@jterc.or.jp.

Transportation Research Record: Journal of the Transportation Research Board, No. 2159, Transportation Research Board of the National Academies, Washington, D.C., 2010, pp. 1–8.
DOI: 10.3141/2159-01

The U.S. government is now evaluating HSR programs to determine which ones should be allocated federal funds. Apart from appraisals of risk mitigation, such as fiscal and institutional capacity, financial plans, commitments from key stakeholders, and adequate project management, the vision indicates the public benefits of an HSR program: (*a*) its contribution to economic recovery efforts, (*b*) its advancement of strategic transportation goals, and (*c*) its furthering of other passenger rail goals articulated in recently passed authorizing legislation (*1*).

The extent to which the U.S. HSR project can realize these public benefits will depend on the location and demand, as well as, perhaps, a strong behavioral change in travel choice mode by the public. However, the overall impact of HSR to society can be confirmed by a detailed examination of HSR projects overseas, which will be useful for the promotion of an efficient and effective HSR system in the United States. For this purpose, this paper examines the findings in the recently published report of the postassessment of Japan's newest HSR routes from the viewpoint of the criteria of public benefits outlined above and explores the implications of those findings on the U.S. HSR project.

OVERVIEW OF THE KYUSHU SHINKANSEN PROJECT

Project Summary and Postassessment

The whole Kyushu Shinkansen Route (the Kagoshima Route) in Japan runs approximately 260 km (161.5 mi) and connects the Hakata and Kagoshima Chuo Stations. (Kyushu Shinkansen consists of the Kagoshima Route and the West Kyushu Route. In this paper, the Kyushu Shinkansen Kagoshima Route is referred to as Kyushu Shinkansen.) The southern half of the segment, which runs 127.6 km (79.3 mi) between Yatsushiro City in Kumamoto Prefecture and Kagoshima City in Kagoshima Prefecture, opened in March 2004 (Figure 1 and Table 1). Although the northern segment is connected to the existing Shinkansen trunk line, the southern segment was given construction priority, because the impact of reducing the travel time as a result of the construction of Shinkansen was estimated to be greater in the southern segment. As a result, unlike Hokuriku Shinkansen and Tohoku Shinkansen, which are connected to the existing Shinkansen Route network, the initial construction of Kyushu Shinkansen did not connect Kyushu Shinkansen to the existing Shinkansen network. Kyushu Shinkansen is also unique, in that transfer from Shinkansen to the limited express, which runs along the conventional train lines at Shin-Yatsushiro, can be done directly on the same platform; and

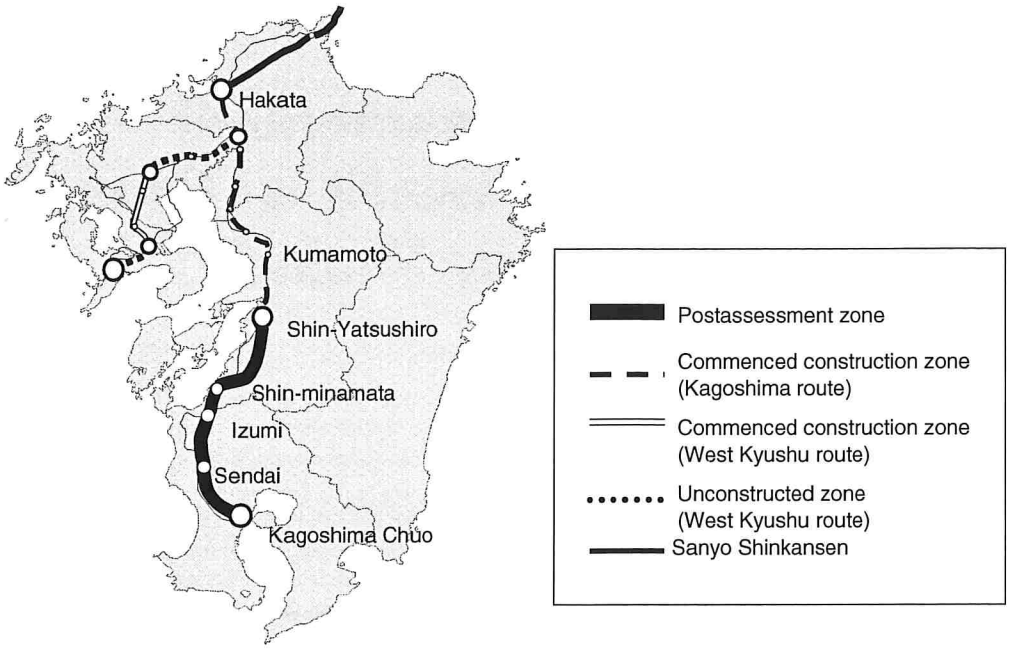


FIGURE 1 Outline of Kyushu Shinkansen (Between Shin-Yatsushiro and Kagoshima Chuo) (2).

the train traffic control, seat reservation, and ticketing systems are integrated for both Shinkansen and the conventional train lines.

With the completion of the route between Shin-Yatsushiro and Kagoshima Chuo, the travel time between Hakata and Kagoshima Chuo was reduced by approximately 90 min, from 3 h 40 min to 2 h 12 min.

Five years have passed since Kyushu Shinkansen started operation, and in fiscal year (FY) 2009 the Japan Railway Construction, Transport, and Technology Agency (JRTT) carried out its postassessment (2), in accordance with the Operating Procedures for Postassessment developed by JRTT. In this paper, which takes into account the relevance of the findings of the postassessment to the public benefits stated in President Obama's vision, the particular focus is on the following, which will be discussed in connection with the vision:

1. Changes in the numbers of passengers carried,
2. The initial effects of the project, and
3. The change in the situation caused by implementation of the project.

This paper introduces for the first time the contents of this postassessment outside of Japan.

Features and Measures Taken at Shin-Yatsushiro Station

Kyushu Shinkansen was built on the standard-gage tracks used for Shinkansen, which are wider than those used for express trains. Therefore, it became necessary for passengers traveling from Hakata to Kagoshima to transfer from the conventional limited express train to Shinkansen at the Shin-Yatsushiro Station. Making a transfer, which is said to take approximately 30 min, significantly influences the convenience of travel. The postassessment describes several measures

TABLE 1 Outline of Project (2)

Item	Summary
Route length	126.1 km
Station names	Kumamoto prefecture: Shin-Yatsushiro station (newly built station), Shin-Minamata station (newly built station) Kagoshima prefecture: Izumi station, Sendai station, Kagoshima Chuo
Construction criteria	Maximum design speed: 260 km/h minimum curve radius: 4,000 m (in principle) Maximum gradient: 35‰ track spacing: 4.3 m Electrification system: 25 kV AC
Types and length of structures	Construction distance: 127.6 km Roadbed: 15.0 km (12%) Bridge: 9.0 km (7%) Viaduct: 15.7 km (12%) Tunnel: 87.9% (69%)

FIGURE 2 Transfer at Shin-Yatsushiro Station (2).

taken at the Shin-Yatsushiro Station to minimize the inconvenience due to transfers.

Same Platform for Transfers

The same platform is used for passengers to transfer from Shinkansen to conventional express trains and vice versa, thereby eliminating the vertical movement of passengers and improving the ease of transfer (Figure 2). The transfer time is just 3 min, and passengers are only required to move across the platform and do need to hurry.

Integrated Services for Shinkansen and Conventional Lines

For all Shinkansen services arriving at the Shin-Yatsushiro Station, connections have been established so that passengers may transfer to conventional limited express trains, and both Shinkansen and the connected limited express trains have the same number of cars. The same applies to the transfer from conventional limited express trains to Shinkansen.

Seat reservations are made in such a way that passengers can be seated in the same seat and carriage of both the Shinkansen and the conventional express trains, thus minimizing the distances that passengers must walk to make a transfer.

When passengers ride both Shinkansen and the conventional limited express trains, a set of three tickets is generally issued and includes a ticket for the distance charge, the Shinkansen train special express ticket, and the conventional limited express train ticket. Only a single combined ticket is issued for travel on Kyushu Shinkansen.

CHANGES IN TRAFFIC VOLUME

The postassessment evaluated the numbers of passengers using each transportation mode and compared the projected number of passengers on Shinkansen and the actual number. These are described here.

Changes in Number of Passengers on Shinkansen

In 2003, before the commencement of Kyushu Shinkansen, the passenger volume was 3,900 per day. After the first year of Kyushu Shinkansen (2004), the actual volume was 8,800 passengers per day (a 225% increase). In the second and third years, the volume of passengers carried was 9,200 per day, and a record high of 9,400 passengers per day was set in the fourth year.

Comparison of the actual performance of transportation density (excluding passengers who used train passes) in FY 2005 with the predicted performance (the value predicted for FY 2000 by the Review Committee on Construction Priority Order organized by the Japanese government and the ruling political party, hereafter referred to as the "predicted value") indicates that the actual performance was 7,400 passenger kilometers per day per kilometer, whereas the predicted value was 8,400 passenger kilometers per day per kilometer. Because the predicted values are for FYs 1995 and 2000, which differ from the year of the actual commencement of Shinkansen (FY 2004), the figures for actual performance used were those for the third year (FY 2005), when the craze over the opening of Shinkansen was considered to have calmed down.

Although there is a gap of approximately 12% between the actual performance in terms of the volume of passengers carried and the predicted value, as stated above, there are differences in other factors, such as population, the precondition of the demand forecast, the economic growth rate, and the transportation service conditions between the time of the prediction and the time of commencement of Shinkansen. A summary of the differences analyzed follows [details are provided elsewhere (*2*)]:

- Population. The actual populations of Kagoshima, Fukuoka, and Kumamoto Prefectures were about 2% to 7% less than the number predicted in 1988.
- Economic growth rate. The actual growth rate was 1% to 4% lower than the value predicted in 1988.
- Level of railway services. The actual travel time was 4 min longer than predicted and a transfer to another train is actually required, which was not expected in 1988. There are some also differences between the actual fares and the predicted values.
- Level of aviation services. The actual travel time is 8 min shorter than predicted. The actual fare is ¥2,320 ($24.4, in July 2009, when ¥95 was equal to $1) higher than the predicted fare.

Sensitivity Analysis of Passengers Carried by Shinkansen: Comparison of Projected and Actual Numbers

To analyze the factors involved with the differences between the actual and the predicted numbers of passengers carried, it was predicted that Kyushu Shinkansen (between the Shin-Yatsushiro and the Kagoshima Chuo Stations) would commence in 2000. A demand forecast was then conducted for different socioeconomic conditions according to the population size, the economic growth rate, and different levels of transportation services; and a sensitivity analysis of demand (traffic volume) for each socioeconomic condition was carried out.

If it is assumed that the population is the predicted value and that the economic growth rate is the same as the actual figure, the result of the estimation of the traffic volume (the transportation density between Shin-Yatsushiro and Kagoshima Chuo) would increase by 6.5%. If it is assumed that the economic growth rate is the predicted value and the population is expected to achieve the actual level, the traffic volume would increase by 6.9%. In the case in which both the population and the economic growth rate are assumed to be the predicted values, the predicted traffic volume would increase by 13.2%.

Sensitivity analysis of the level of transportation services was conducted for the origins–destinations (O-Ds) between the Fukuoka

City and Kagoshima City zones by changing such conditions as the travel time and fares. If the railway travel time is reduced by 4 min, which is the predicted value, the impact on demand would increase by 1.5%. It increased by 3.4% if it was supposed that there was no transfer to another train at the Shin-Yatsushiro Station and by 2.2% if the railway fare was assumed to be the same as the predicted condition. When air travel times and fares changed, the demand for Shinkansen decreased by 1.7%. If all the conditions described above were changed, demand increased by 5.4%.

The postassessment report indicates that on the basis of the analysis presented above, the differences between the predicted demand and the actual demand for Kyushu Shinkansen were affected to a greater extent by changes in socioeconomic conditions than other factors.

Changes in Passenger Use of Each Mode

The postassessment evaluated the number of passengers using each transportation mode between Fukuoka and Kagoshima Prefectures, as shown in Figure 3.

Evaluation of the changes in the numbers of passengers carried by railway showed that the volume had been stable before the commencement of the Shinkansen operation but significantly increased in 2004. Evaluation of the trend in actual aviation usage showed that the numbers of passengers started to show a sharp decline in FY 2003, a year before Kyushu Shinkansen commenced, and dropped to 250,000 a year in FY 2006, which is about 35% of the number of passengers at the peak. The share of transportation by air was 43% in FY 1990 but dropped to 12% in FY 2006. (It should be noted that the number of air passengers declined in 2003 because Air Nippon suspended four round-trip services a day because of low profits on the route.)

Since 1991, the volume of users of the highway express bus that runs between Fukuoka and Kagoshima Prefectures had been on the increase, especially by the commencement of Shinkansen, and the increased volume was considered to have a positive impact on the image of the area along Shinkansen. As the volume of highway bus users between Kumamoto and Kagoshima Prefectures had been flat since 1990, the service between those two locations was discontinued with the commencement of Kyushu Shinkansen, whereas a new route between Hitoyoshi and the Kagoshima Airport was opened.

According to the Road Traffic Census compiled by the Ministry of Land, Infrastructure, Transport, and Tourism in Japan (MLIT), it was observed for the segments studied that the number of vehicles traveling on the Kyushu Expressway did not change much after the commencement of Kyushu Shinkansen but did increase slightly in all zones.

KYUSHU SHINKANSEN AND PUBLIC BENEFITS

President Obama proposed the highest level of funding that the United States has ever provided for HSR. In the president's vision, the U.S. Department of Transportation indicates that evaluation of the funding for the program should be based on transparent, merit-based selection criteria of public benefits and risk mitigation (*1*). Although risk mitigation, such as the capacity to carry out projects and financial plans, can be a general prerequisite of a project in any field and its evaluation is beyond the scope of this paper, the level of public benefits provided by HSR can vary across projects. This section examines the postassessment of Kyushu Shinkansen from the viewpoint of the public benefits described in the vision.

Effects on Economy

The public benefits in the vision are explained as "the extent to which the project or corridor program provides specific, measureable, achievable benefits in a timely and cost-effective manner" (*1*). The first item among these public benefits is economic recovery. The postassessment of Kyushu Shinkansen confirmed the benefits

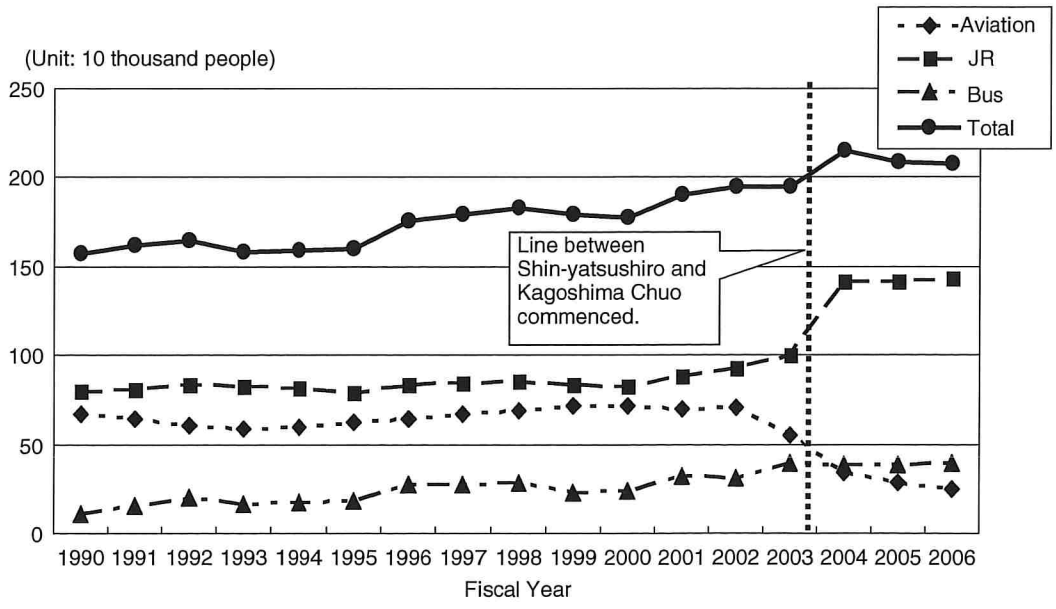


FIGURE 3 Numbers of passengers by each mode of transportation between Fukuoka and Kagoshima prefectures. [Source: JRTT (*2*), based on Passenger Movement Survey in On-Line Zone, compiled by MLIT.]

by use of a cost–benefit analysis and showed that the project had effects that rippled through the economy.

Cost–Benefit Analysis

Project efficiency is evaluated from a socioeconomic point of view by cost–benefit analysis. The benefit of the project is assessed by computation of the immediate benefit by totaling the user and supplier benefits, in accordance with the *Assessment Guidance Manual 2005 for Railroad Project,* edited by MLIT (*3*). The costs include those for construction, the expense for trains, and maintenance costs.

Table 2 shows the actual and predicted demands used in the analysis of the postassessment. Other conditions and the model of demand prediction are explained by JRTT (*2*).

The cost–benefit analysis over the 50 years after the commencement of the service was conducted on the basis of the assumption that the Shinkansen Route, between Hakata and Shin-Yatsushiro, is not yet open (it is scheduled to open in 2010). The computational results of the analysis are as follows (the amount of the benefit and the cost is a total amount that is converted to the present value and accumulates for 50 years since the time of operational commencement):

benefit = ¥1,030.4 billion ($10.8 billion converted at ¥95 to $1 in July 2009),
cost = ¥913.9 billion ($9.6 billion at the same exchange rate),
net present value = ¥116.5 billion ($1.2 billion at the same exchange rate), and
cost–benefit ratio = approximately 1.1.

The travel time reduction in the zone between the Hakata and the Kagoshima Chuo Stations was large (approximately 90 min). However, the analysis was done for a zone with a relatively small demand (between the Shin-Yatsushiro and Kagoshima Chuo Stations) and was based on the assumption that service in the zone with greater demand (between Hakata and Shin-Yatsushiro) had not commenced; thus, the cost–benefit ratio remained rather small. Note that the cost–benefit ratio calculated in the postassessment for Hokuriku Shinkansen (Takasaki–Nagano), which has operated since October 1997, was 1.8 (*4*) and that that for Tohoku Shinkansen (Morioka–Hachinohe), which has operated since December 2002, was 1.3 (*5*). If the analysis was based on the assumption that the route between Hakata and Shin-Yatsushiro had commenced, the cost–benefit ratio should be bigger than the ratio given above.

Impact on Overall Economy

The construction of Kyushu Shinkansen led to improvements in the speed and frequency of transportation services. These improvements brought about the streamlining of business activities, for instance, reductions in expenses for business trips, the expansion of business opportunities, efficient means of information gathering, and easier means of holding meetings and business negotiations. Communication costs go down as interregional proximity rises, and an economic ripple effect occurs as productivity improves.

Calculation of the economic ripple effect in the postassessment, based on the production function method, shows that it provided an annual increase in the gross domestic product in 2008 (the fifth year) of approximately ¥25 billion ($263 million, converted at ¥95 to $1; 0.005% in the gross domestic product), that it will be approximately ¥29 billion ($305 million at the same exchange rate) in 2013 (the 10th year), and that it will have the same effect of increasing the gross domestic product by 0.005%.

Strategic Transportation Goals

After it addresses economic recovery efforts, the vision addresses strategic transportation goals that will benefit the public (*1*). It is said that a transportation investment strategy needs to meet several goals: (*a*) to ensure the availability of safe and efficient transportation choices, (*b*) to build a foundation for economic competitiveness, (*c*) to promote energy efficiency and environmental quality, and (*d*) to support interconnected, livable communities.

Safety

The vision encourages the promotion of the safest possible movement of goods and people (*1*). In this context, Shinkansen is a means of transportation with a strong reputation for safety, as there has been no fatal accident among any of the passengers of Shinkansen. The safety benefits are massive when the death tolls from the use of automobiles and conventional railways are considered.

Efficiency

Along with safety, the vision urges optimization of the use of the existing and new transportation infrastructure (*1*). In this regard, Kyushu Shinkansen contributed to reducing the travel time while making it possible to stay at the destination for a longer period of time.

Reduced Travel Time Shinkansen reduced the travel time between Hakata and Kagoshima by approximately 90 min. Compared with the travel time by air over the same route, the travel time of the Shinkansen is even 10 min shorter.

In addition, the travel time by rail between Kumamoto and Kagoshima was shortened by more than half, from 2 h 23 min before

TABLE 2 Demand for Transportation Used in Cost–Benefit Analysis (*2*) (noncommuter pass, passenger kilometers per day per kilometer)

	Actual Performance	Predicted Demand			
	FY 07	FY 10	FY 20	FY 30	FY 50
Case 1. New Shinkansen line not constructed	—	4,000	4,000	3,900	3,400
Case 2. New Shinkansen line constructed	8,600	8,300	8,200	8,100	6,900

SOURCE: JRTT (*2*).

the commencement Shinkansen to 58 min after. However, as the highway express bus service between Kumamoto and Kagoshima was terminated in March 2004, when Shinkansen opened, users of the highway express bus from Kumamoto to Kagoshima or vice versa must now change buses.

Increase in Time Available at Destination for Day Trips The time available at each destination for day trips increased significantly. In the case of the Kagoshima Chuo Station, for those who left Hakata by the first train, the time available at the destination increased substantially (approximately 4 h 10 min), from approximately 8 h 30 min to approximately 12 h 40 min. The time available at the destination at the Hakata Station for those who left the Kagoshima Chuo Station also showed a large increase (approximately 2 h 40 min).

The time available at the destination at the Kagoshima Chuo Station for those who left the Kumamoto Station by the first train increased a great deal (approximately 6 h), from approximately 8 h 30 min to approximately 14 h 30 min, because, in addition to the reduction of travel time by Shinkansen, an early-morning first train service leaving Kumamoto and a relaying limited express service (Relay Tsubame) leaving for Kumamoto at a late hour were newly launched. The time available at the Kumamoto Station for those who left the Kagoshima Chuo Station by the first train also increased (approximately 2 h 50 min).

Economic Competitiveness

The Obama Administration calls for a new approach to transportation to build a foundation for economic competitiveness. It is expected to stimulate economic growth by facilitating the efficient movement of people and goods. The Kyushu Shinkansen example supports HSR by showing that it provides newly created travel demand, enhances business activities, and creates a surge in tourism.

Creation of New Travel Demand According to the responses to the Kyushu Shinkansen user questionnaire to a question on whether the responder would have gone on a trip if Shinkansen had not been commenced, 17.8% responded that they would not have gone on a trip or that they would have changed the destination. This indicates that the operational start of Kyushu Shinkansen created new travel demand. The number of Kyushu Shinkansen users in 2004 was approximately 3.5 million per year. If the demand created is assumed to be 17.8%, 310,000 travelers can be considered to be newly generated per year.

Effects on Business and Commercial Activities According to the questionnaire survey on Kyushu Shinkansen conducted by the Kyushu Economic Research Center, which has branch offices and regional companies in the prefectures on the Kagoshima Honsen Line, approximately 72% of the companies located in Kagoshima Prefecture, 64% of those located in Fukuoka Prefecture, and 22% of those located in Kumamoto Prefecture responded that the commencement of Kyushu Shinkansen had a positive impact on their businesses. Meanwhile, the proportion that responded that it had a negative impact was less than 5%. A typical impact includes reduced expenses for business trips, the expansion of business operations, and easier information gathering.

Other positive impacts on business include an increased number of conventions being held in Kagoshima City and increased numbers of participants at those conventions, which shows that the construction of Shinkansen contributed to greater business exchange.

Changes in Number of Inbound Tourists The tourism statistics compiled by local governments show that the number of tourists inbound to the regions along Kyushu Shinkansen in 2004, when service commenced, increased over that in the previous year. Although the operational start of Shinkansen is considered to have contributed to this increase, the trend since FY 2005 has varied by region and has possibly been influenced by economic trends, weather, and other factors. Thus, the direct impact of Shinkansen is not necessarily clear. Mid- to long-term analyses will be required to determine the true direct impact.

Trends in Overnight Guests According to the tourism statistics prepared by Kagoshima Prefecture, the numbers of overnight guests to Kagoshima Prefecture from outside the prefecture were approximately 7.6 million in 2003, the year before the commencement of Kyushu Shinkansen; approximately 7.8 million in 2004, the year of commencement; and approximately 7.9 million in 2007, a record high. In FY 2007, the numbers of overnight guests from outside the prefectures increased approximately 115% in the southern regions of Kyushu (Kumamoto, where Shinkansen operates, and Miyazaki, which is adjacent to Kumamoto) and approximately 110% in the northern and middle regions of Kyushu, where no Shinkansen is operating. These percentages are relative to the numbers in FY 2003. It seems that factors such as improved convenience and improved image were more easily conveyed in the areas close to Kyushu Shinkansen.

No tendency was seen for inbound guests staying overnight to have converted to being 1-day-trip tourists because of the commencement of Shinkansen; rather, the ratio of overnight guests increased. Analysis of the area of accommodation on the basis of the Kyushu Shinkansen-Related Tourism Trend Survey (March 2005), prepared by Kagoshima Prefecture, showed that approximately 75% of those visiting the Kagoshima region stayed in areas other than Kagoshima. Thus, it is fully possible that more tourists are visiting the areas along Shinkansen and that more tourists are going to other areas not along Shinkansen.

Energy Efficiency and Environmental Quality

The U.S. HSR projects emphasize that HSR is green by claiming that today's intercity passenger rail service consumes one-third less energy per passenger mile than cars. The strategic transportation goals also require energy independence and renewable energy to be fostered and the amounts of pollutants and greenhouse gas emissions to be reduced. The postassessment of Kyushu Shinkansen analyzed the environmental effects.

The CO_2 emissions caused by transportation mode in the zones between Fukuoka and Kagoshima Prefectures and between Kumamoto and Kagoshima Prefectures before and after the commencement of Shinkansen were calculated on the basis of the number of interprefectural O-Ds presented in the net movement survey of mainline passengers and are shown in Figure 4. Note that the number of interprefectural O-Ds of private vehicles in the FY 2005 survey declined, but it seems inappropriate to assume that the traffic was shifted to railways. Therefore, the CO_2 emissions from private vehicles are excluded from the analysis of the postassessment. The details of the environmental impacts are explained elsewhere (2).

The CO_2 emission from railways increased because the number of users increased after Shinkansen started operation, whereas that due to aviation, which has higher levels of basic emissions than railways, declined. The total CO_2 emissions of all transportation modes in the zones between Fukuoka and Kagoshima Prefectures and between

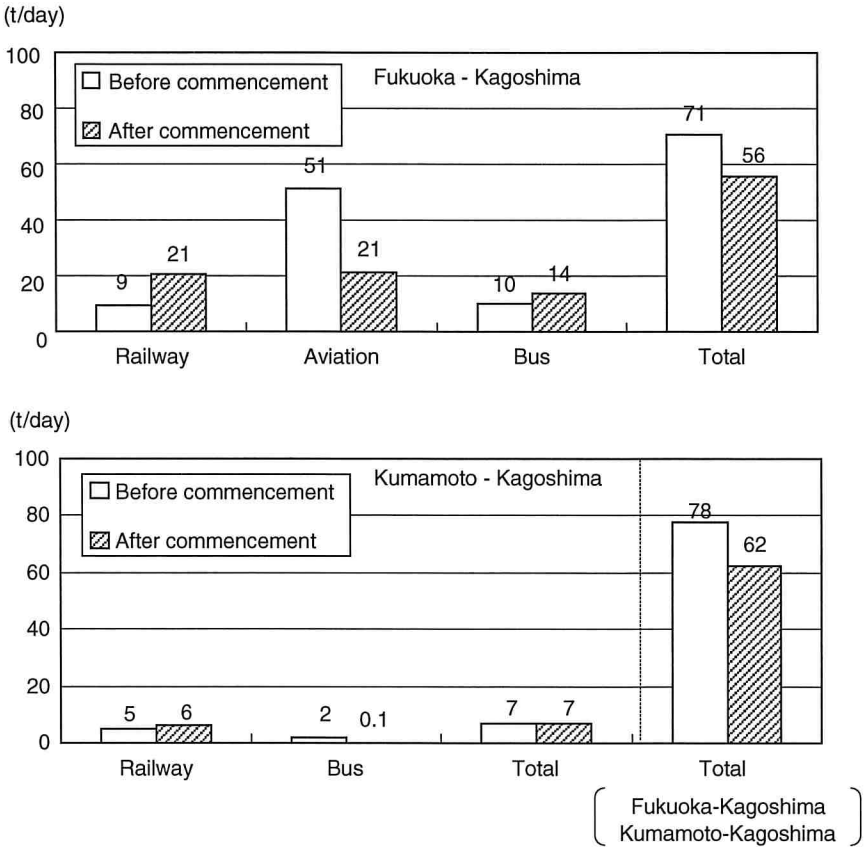


FIGURE 4 CO_2 emissions by different modes of transportation before and after commencement of the Kyushu Shinkansen operation. [Source: JRTT (2), based on number of interprefectural ODs presented in National Net Movement Survey of Main-Line Passengers (FY 2000 and FY 2005) by MLIT.]

Kumamoto and Kagoshima Prefectures decreased to approximately 80% of the level of emissions before the commencement of Kyushu Shinkansen, from 78 to 62 metric tons per day. Therefore, the construction of Shinkansen is considered to contribute to the reduction in the overall environmental load.

Effects on Livable Communities

Vice President Joseph Biden said on April 16, 2009, when the vision was released, "Everyone knows railways are the best way to connect communities to each other," and investing in an HSR system will "loosen the congestion suffocating our highways and skyways." The last part of the strategic transportation goals concerns livable communities, in which quality of life in local communities is improved by promoting affordable, convenient, and sustainable housing, energy, and transportation options. The postassessment of Kyushu Shinkansen reveals that the public chose Shinkansen because of its convenience for traveling and commuting. Shinkansen is thus contributing to local communities.

Transport Share and Changes in Use of Transportation Modes Since the commencement of Kyushu Shinkansen in March 2004, the share of travel by railway for the zone between Fukuoka and Kagoshima Prefectures increased from 41% to 71% and the share of travel by air decreased from 42% to 12%. Although the number of bus users increased, the total share of transportation by bus fell from 18% to 17%, because the number of railway users increased substantially.

As for the transportation share in the zone between Kumamoto and Kagoshima Prefectures, the share of travel by railway rose from 88% to 99.5%, whereas the share by bus fell from 12% to 0.5% because of the termination of the highway express bus services.

The questionnaire survey was targeted at Shinkansen users and was conducted on the basis of the mode of transportation that the participants used before the operational start of Shinkansen. The findings indicated that 20% of all Shinkansen users changed from air to Shinkansen and 25% switched from driving a car to riding Shinkansen.

Evaluation of the purpose of the trip showed that 33% of the Shinkansen users for business switched from air to Shinkansen, whereas approximately 35% of users traveling for leisure and recreation changed from the automobile to the Shinkansen service.

Impact on Work and School Commuting As the travel time was reduced on the commencement of Kyushu Shinkansen, the level of use of the railway for work and school commuting from Izumi City and Satsuma Sendai City to Kagoshima City and other cities increased significantly. In the second year of the Kyushu Shinkansen service as well as later, the number of commuters increased. In the third year (as of January 31, 2007), the number of commuters using the railway was 1,100 a day, which is approximately 11 times that before the launch of Shinkansen (approximately 100 as of January 31, 2004).

IMPLICATIONS OF HSR PROJECT FOR THE UNITED STATES

The vision for HSR illustrates the fact that the comparative advantage of each transportation mode varies according to market factors and that HSR can play a critical role, particularly in 100- to 600-mi intercity corridors. It is also possible that if the speed of U.S. railways improves to the speeds of Japanese and European counterparts, this distance could be even longer, as many as 800 mi, for example.

Japan's case of Kyushu Shinkansen has significant implications for the U.S. HSR plan. First, the economic benefits of HSR have been confirmed. The Kyushu Route, which opened in March 2004, is only a part of the overall planned route but has already generated significant demand and widened the area of economic activity. The members of the public have actually changed their travel behaviors. The impact is expected to increase even more when operations on the entire route begin.

Second, without further study of other transportation modes, it is still not clear that these benefits of HSR could outweigh those of other modes. However, among the public benefits cited in the vision, the environmental benefit of HSR is worth noting, particularly in the context of climate change. HSR is literally a green form of transportation compared with the air and private vehicle modes of transportation, as Table 3 shows. Given the comparative market advantage of HSR for trips between 100 and 600 mi, the potential environmental benefits of HSR are high if air travel shifts to railways. The Institution for Transport Policy Studies (ITPS) has recently conducted a study that illustrates that even if electric vehicles are fully disseminated and the rate of use of biofuels around the world is high, without much of a reduction in the levels of CO_2 emissions from aircraft, by 2050 the total global CO_2 emissions from the transportation sector will not decrease much compared with those in 2000 [ITPS, unpublished data, based on a simulation with the IEA/SMP transportation model, World Business Council for Sustainable Development (*6*); the IEA/SMP transport model is a spreadsheet-based global transportation model developed by the International Energy Agency and the World Business Council for Sustainable Development's Sustainable Mobility Project]. This indicates that the key factor in the long-term reduction of CO_2 levels is how air passengers can be shifted to other, more environmentally friendly modes. The postassessment of Kyushu Shinkansen demonstrates that HSR may contribute to this by shifting passengers from air to rail travel. To gain environmental benefits large enough to contribute to the whole transportation sector, however, the experience with Kyushu Shinkansen also indicates that a certain scale of railway infrastructure is necessary.

TABLE 3 CO_2 Emission and Consumption Rates

	Railways	Aviation	Bus	Private Automobile
g-CO_2/passenger kilo	18	111	51	172
In case consumption rate of railroad equals 1.0	1.0	6.2	2.8	9.6

NOTE: The figures in *Transportation, Traffic and Environment 2008* reflect the numbers in FY 06.
SOURCE: Prepared by JRTT (*2*) based on *Transportation, Traffic and Environment 2008* by Foundation for Promoting Personal Mobility and Ecological Transportation.

Third, the Kyushu example gives good lessons for the actual HSR plan in terms of how the HSR can connect to the conventional line, such as in the case of California. If HSR and conventional Amtrak services are connected on the same platform so that passengers can easily transfer from one service to another, such a system can increase demand for travel by rail. One should also be reminded that the economic success of Kyushu Shinkansen is partly derived from the order of construction, in which construction in the area with the greatest benefits from reduced travel times was prioritized. If the partial speedup of California railway construction is considered, the priority for construction should be the area that will provide the greatest benefits. Moreover, having seen the effects on commuters in Kyushu, if the speed of the commuter line is improved, an increase in commuting demand can be expected.

CONCLUSION

The postassessment of Kyushu Shinkansen by JRTT confirms the increase in the numbers of passengers with the commencement of the new Shinkansen Route. Although the degree of the increase was smaller than expected, the sensitivity analysis indicated that the smaller-than-expected increase was due to changes in the socioeconomic situation. The results of the postassessment support the U.S. HSR strategic plan by showing the economic effects as well as other public benefits of HSR, including its contribution to the environment.

However, to maximize the gains achieved by the implementation of HSR, the postassessment also implies that large-scale HSR infrastructure is needed. The vision for HSR in America is just a starting point. The United States needs to extend its funding for HSR so that the public can better understand its usefulness and more fully enjoy its economic and environmental benefits. Moreover, evaluation of the shift of traffic from private cars to railways by use of the Kyushu example is limited. Further study is required to explore changes in travel behavior.

REFERENCES

1. U.S. Department of Transportation. *Vision for High-Speed Rail in America.* www.fra.dot.gov/Downloads/RRDev/hsrstrategicplan.pdf. Accessed April 16, 2009.
2. *Report on Post-Assessment of Kyushu Shinkansen (between Shin-Yatsushiro and Kagoshima Chuo)* (in Japanese). Japan Railway Construction, Transport, and Technology Agency, Yokohama, Japan, March 2009. www.jrtt.go.jp/info/data/jk20-5-2.pdf. Accessed May 15, 2009.
3. Ministry of Land, Infrastructure, Transport, and Tourism. *Assessment Guidance Manual 2005 for Railroad Project* (in Japanese), June 2005. www.mlit.go.jp/tetudo/jigyo_hyoka/manual2005.html. Accessed May 15, 2009.
4. *Report on Post-Assessment of Hokuriku Shinkansen (between Takasaki and Nagano)* (in Japanese). Japan Railway Construction, Transport, and Technology Agency, Yokohama, Japan, March 2008. www.jrtt.go.jp/info/data/jk19-5-2.pdf. Accessed May 15, 2009.
5. *Report on Post-Assessment of Tohoku Shinkansen (between Morioka and Hachinohe)* (in Japanese). Japan Railway Construction, Transport, and Technology Agency, Yokohama, Japan, March 2008. http://www.jrtt.go.jp/info/data/jk19-6-2.pdf. Accessed May 15, 2009.
6. *Global Transport Spreadsheet Model.* World Business Council for Sustainable Development, Geneva, Switzerland. http://www.wbcsd.org/web/publications/mobility/smp-model-spreadsheet.xls. Accessed April 20, 2009.

The Intercity Passenger Rail Committee peer-reviewed this paper.

High-Speed Rail Project Development Processes in the United States and China

Xueming (Jimmy) Chen and Ming Zhang

Amid increasing global economic competition, both the United States and China are investing in high-speed rail (HSR) as a national strategy to enhance their competitiveness. China is rapidly advancing by laying down HSR tracks, while the United States is deliberating the environmental, social, and economic implications of HSR development before embarking on actual construction. With the California High-Speed Rail Project and the Beijing–Shanghai Express Railway Project as case studies, this paper juxtaposes the HSR development processes concerning investment decisions, planning, and implementation in the United States and China and draws lessons informative to strategic infrastructure investments in both countries.

Increasing global competition has stimulated many countries to develop high-speed rail (HSR) as a key strategy to enhance the national transportation infrastructure and to strengthen economic competitiveness. HSR has impressive track records in European countries and advanced Asian countries, such as Japan and South Korea. Recently, China has been catching up in laying down HSR tracks. Poised to become the world's leader in HSR, China will build 42 high-speed passenger rail lines with a total length of 13,000 km by the year 2012 (*1*). For reference, as of early 2008, there were approximately 10,000 km of new HSR lines in operation in the world (*2*).

Comparatively, the United States has lagged behind in rail development. From 1978 to 1999, federal spending on rail transportation amounted to only $18.3 billion, 3.6% of total national spending in transportation (*3*). The federal government established the National Railroad Passenger Corporation (Amtrak) in 1971 to maintain rail passenger services and invested billions of dollars in the Northeast Corridor Improvement Project. Still, rail has been precipitously losing its market share, and the share of passenger rail at the federal level in intercity transportation investment continues to shrink (*4*). Wolman noted that long distances between major cities, cheap gasoline prices, and advanced highway and aviation systems are among the most important reasons why the United States has not made any major investment in HSR (*5*). It has also been noted that challenging physical constraints, demographics, and institutions not being suited to central solutions, along with resistance to centralized planning, are believed to hamper HSR progress in the United States (*6*).

The year 2009 witnessed a major turn of interest to rail transportation in the United States, when President Barack Obama unveiled his administration's blueprint for a new national network of HSR lines. Specifically, the blueprint under the American Recovery and Reinvestment Act of 2009 (ARRA) commits $8 billion as a down payment to the development of HSR corridors, along with $1 billion per year for at least 5 years (proposed in the fiscal year 2010 budget). In addition, ARRA commits $1.3 billion for capital improvements to Amtrak, including on the Northeast Corridor. However, as recognized by FRA, mounting challenges in advancing this new passenger rail vision lay ahead, ranging from federal and state fiscal constraints, relationships with private railroads, and multistate partnerships to local political, institutional, and business resistance. In addition, many critics still question the wisdom of developing HSR in the United States (*7*).

This paper reviews from an institutional perspective the HSR development processes in the United States and China and presents two specific cases: the developments of the California High-Speed Rail Project (CHSRP) and the Beijing–Shanghai Express Railway Project (BSERP). Recognizing the drastic differences between the two countries in political, economic, cultural, geographical, and other aspects, the paper does not intend to offer a comparative study in a strictly methodological sense. Instead, it juxtaposes the process of HSR investment decisions, planning, and implementation in the two countries and regions. Growing interest in HSR in the two countries and the rest of the world has motivated this study. By juxtaposing the two cases, the paper is aimed at generating inspirational lessons useful to the development of HSR in the respective country and regional settings, although the country- and region-specific experiences may not be transferred directly from one place to another.

After this introduction, the rest of the paper contains four parts. The next section introduces the proposed HSR networks in the United States and China. The paper then presents information on the project technical specifications, project management, planning and design, funding, right-of-way acquisition, construction, and approval processes for CHSRP and BSERP. The similarities and differences in HSR development illustrated in the two cases are then traced to each country's institutional characteristics. Finally, concluding remarks are offered and implications for HSR development in each of the two countries are drawn.

Before this case study is unfolded, two caveats are given. First, it is worth noting that there are a number of subtle differences in defining HSR in the United States and China. China follows the International Union of Railways by defining HSR as a type of passenger rail transportation that operates at speeds of 200 km/h and faster. In contrast, FRA distinguishes three types of HSR: (*a*) express HSR, which is frequent, express service between major population centers

X. Chen, L. Douglas Wilder School of Government and Public Affairs, Virginia Commonwealth University, Scherer Hall 511, 923 West Franklin Street, Richmond, VA 23284. M. Zhang, School of Architecture, University of Texas at Austin, 1 University Station B7500, Austin, TX 78712. Corresponding author: X. Chen, xchen2@vcu.edu.

Transportation Research Record: Journal of the Transportation Research Board, No. 2159, Transportation Research Board of the National Academies, Washington, D.C., 2010, pp. 9–17.
DOI: 10.3141/2159-02

200 to 600 mi (320 to 965 km) apart and which achieves top speeds of at least 150 mph (240 km/h) on completely grade-separated, dedicated rights-of-way; (*b*) regional HSR, which is relatively frequent service between major and moderate population centers 100 to 500 mi (160 to 800 km) apart and which achieves top speeds of 110 to 150 mph (177 to 240 km/h), which is grade separated, and which has dedicated and shared tracks; and (*c*) emerging HSR, which is the development of corridors of 100 to 500 mi (160 to 800 km) and which has top speeds of up to 90 to 110 mph (145 to 177 km/h) primarily on shared tracks (*4*). Second, it is incorrect to simply assume that HSR is superior to other types of long-distance transportation modes under all circumstances. HSR is mostly preferred for travel distances of 200 to 500 mi.

PROPOSED HSR NETWORKS IN THE UNITED STATES AND CHINA

On April 16, 2009, President Barack Obama unveiled his administration's blueprint for a new national network of HSR lines, including 10 high-speed intercity corridors potentially eligible for federal funding. These corridor designations were made on the basis of state applications for corridors in which trains are expected to achieve operating speeds of at least 90 mph to address highway–rail grade crossing issues.

On the Chinese side, the proposed passenger dedicated lines (PDLs) with operating speeds of more than 200 km/h constitute the backbone of the future Chinese HSR network. The China Ministry of Railways (MOR) has designated four major north–south PDLs and four major east–west PDLs. In addition to these eight major PDLs serving long-distance travel, the Chinese HSR network also includes

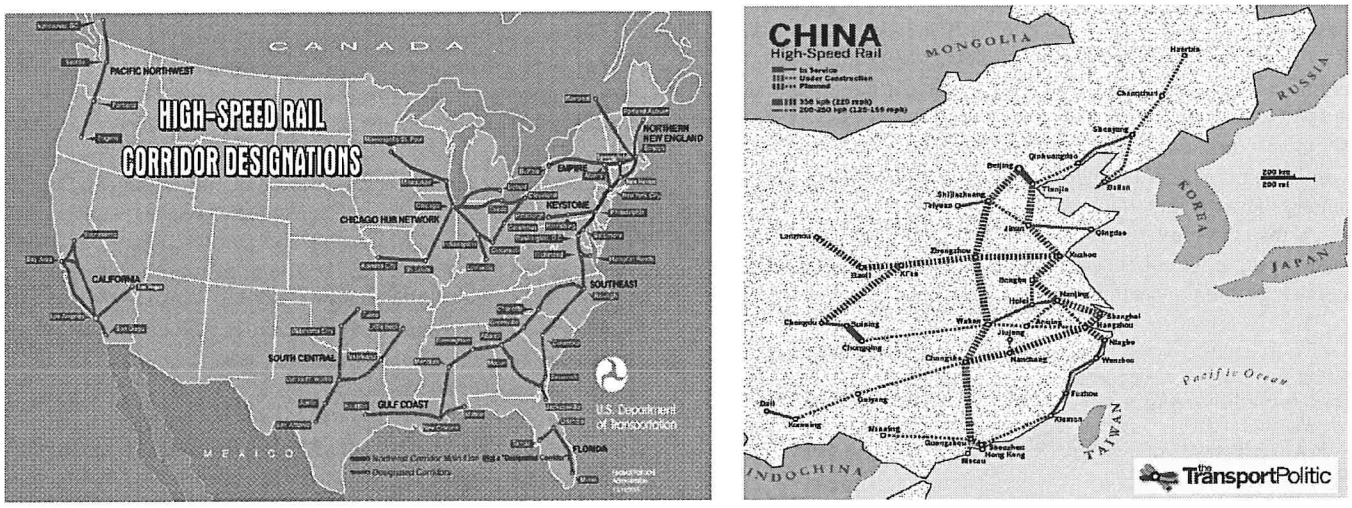
- Intercity lines with maximum speeds of 200 to 250 km/h built in the largest metropolitan areas, for example, between Beijing and Tianjin;
- Upgraded conventional railways, which are those main lines on conventional tracks upgraded so that trains may achieve maximum speeds of 200 km/h; and
- Other similar rail lines, which are those passenger and freight rail lines with standard speeds of 200 to 250 km/h that are not included in the long-term PDL plan.

Figure 1 and Tables 1 and 2 compare the U.S. and Chinese HSR networks (*8–10*). From these lists of HSR projects in the United States and China, this paper selects and presents in further detail two cases, CHSRP and BSERP, primarily because of their comparable project scales and their relatively close project readiness statuses.

PROJECT DESCRIPTIONS OF CHSRP AND BSERP

Once it is built, CHSRP, which will be 1,280 km (800 mi) long, will link the San Francisco and Los Angeles, California, metropolitan areas (Phase 1) in 2.5 h. The planned system would also ultimately serve Sacramento, San Diego, and other cities in California (Phase 2). As one of the most important states in the United States, California has 4.3%, 12.2%, and 13.1% of the country's land, population, and gross domestic product (GDP), respectively. The construction of CHSRP will commence no later than 2012, the federally mandated deadline for receiving funds. It is expected to take 8 to 11 years before the operation of its initial segment. The entire project is planned to be completed by 2030 (*11*).

BSERP is an HSR system 1,318 km (819 mi) long that connects the core areas of China's east coast. This rail corridor traverses the following important provinces and municipalities in China: Beijing, Tianjin, Hebei, Shandong, Anhui, Jiangsu, and Shanghai (Figure 2) (*12, 13*). The corridor provinces and provincial-level municipalities account for 6.5%, 26.7%, and 43.3% of China's land, population, and GDP, respectively. In 2007, the average population density of the corridor region was 556 people per square kilometer, more than six times that of the California Corridor (89 people per square kilometer). Both demographically and economically, BSERP is more important to China than CHSRP is to the United States. Construction of BSERP started on April 18, 2008, and is scheduled to be completed by 2012 (*1*). Figure 2 and Table 3 show the alignments and construction sections of both rail lines, respectively (*14, 15*).



(a) (b)

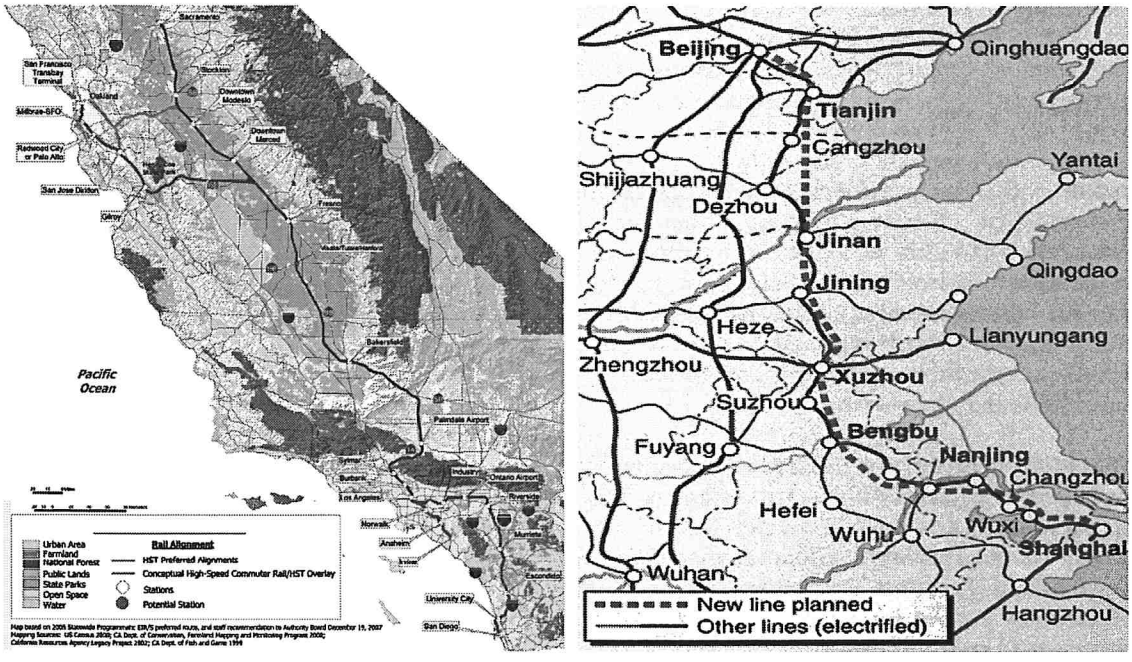
FIGURE 1 HSR networks of (*a*) United States and (*b*) China (*8, 9*).

TABLE 1 U.S. High-Speed Rail Corridors (*8*)

Corridor Name	Major Cities Linked
California Corridor	Sacramento, San Francisco, San Jose, Los Angeles, San Diego
Pacific Northwest Corridor	Eugene, Ore.; Portland, Ore.; Tacoma, Wash.; Seattle, Wash.; Vancouver, B.C.
South Central Corridor	Tulsa, Okla.; Oklahoma City, Okla.; Little Rock, Ark.; Dallas, Tex.; San Antonio, Tex.
Gulf Coast Corridor	Houston, Tex.; New Orleans, La.; Mobile, Ala.
Chicago Hub Network	Chicago, Ill.; Detroit, Mich.; Cleveland, Ohio; Kansas City, Mo.; St. Louis, Mo.; Louisville, Ky.; Milwaukee, Wisc.; Minneapolis–St. Paul, Minn.
Florida Corridor	Tampa, Orlando, Miami
Southeast Corridor	Raleigh, N.C.; Charlotte, N.C.; Atlanta, Ga.; Columbia, S.C.; Jacksonville, Fla.
Keystone Corridor	Pittsburgh, Pa.; Philadelphia, Pa.
Empire Corridor	Buffalo, N.Y.; Albany, N.Y.
Northern New England Corridor	Boston, Mass.; Portland/Auburn, Maine; Montreal, Canada; Springfield, Mass.; New Haven, Conn.

TABLE 2 HSR Corridors in China (*10*)

PDL Direction	PDL Name
East–west	Beijing–Harbin Line via Tianjin, Qinhuangdao, Shenyang. Branch: Shenyang-Dalian Beijing–Shanghai Express Railway Line via Tianjin, Jinan, Xuzhou, Bengbu, Nanjing Beijing–Hong Kong Line via Shijiazhuang, Zhengzhou, Wuhan, Changsha, Guangzhou, Shenzhen Shanghai–Shenzhen Line via Hangzhou, Ningbo, Wenzhou, Fuzhou, Xiamen
North–south	Qingdao–Taiyuan Line via Jinan, Shijiazhuang, Taiyuan Xuzhou–Lanzhou Line via Zhengzhou, Xi'an, Baoji Shanghai–Chengdu Line, via Nanjing, Hefei, Wuhan, Chongqing Hangzhou–Kunming Line via Nanchang, Changsha, Guiyang



(a) (b)

FIGURE 2 Alignments of (*a*) CHSRP and (*b*) BSERP (*12*, *13*).

TABLE 3 CHSRP and BSERP Construction Sections (*14, 15*)

	CHSRP			BSERP		
Section	From	To	Distance (km)	From	To	Distance (km)
1	San Francisco	San Jose	77	Beijing	Tianjin	131
2	San Jose	Merced	201	Tianjin	Jinan	288
3	Merced	Fresno	90	Jinan	Xuzhou	269
4	Fresno	Bakersfield	113	Xuzhou	Suzhou	79
5	Bakersfield	Palmdale	182	Suzhou	Bengbu	77
6	Palmdale	Los Angeles	93	Bengbu	Nanjing	174
7	Los Angeles	Anaheim	47	Nanjing	Wuxi	183
8	Merced	Sacramento	177	Wuxi	Suzhou	26
9	Los Angeles	San Diego	269	Suzhou	Shanghai	75

SOURCE: http://www.cahighspeedrail.ca.gov/library/default.aspx?name=Federal%20Stimulus%20Update; http://en.wikipedia.org/wiki/Beijing-Shanghai_Express_Railway.

Table 4 compares the technical specifications for CHSRP and BSERP (*15, 16*). Notably, the two HSR projects have comparable track lengths, design speeds, and numbers of stations. Because of the much higher population density and the much lower automobile ownership rate in China than in the United States, BSERP is expected to generate a much higher rail ridership than CHSRP. Higher land and labor costs in the United States make CHSRP more than 40% more expensive to construct than BSERP.

CHSRP AND BSERP PROJECT DEVELOPMENT PROCESSES

This section compares the CHSRP and BSERP project development processes, with emphases being placed on project management, planning, funding, and construction.

Project Management and Planning

Established in 1996 by the California State Legislature, the California High-Speed Rail Authority (CHSRA) is the state entity responsible for planning, designing, constructing, and operating a high-speed train system serving California's major metropolitan areas. CHSRA has a nine-member policy board (five members are appointed by the governor, two are appointed by the California Senate Rules Committee, and two are appointed by the speaker of the California Assembly) and a small core staff. These board members have diverse backgrounds (they are governmental officials, attorneys, business leaders, community activists, and others) but have equal power in the decision-making process. The board chair serves on a rotational basis and presides over the board meetings. As a state entity, CHSRA is actually a coalition-building organization intended to build HSR in the Golden State.

TABLE 4 Technical Specifications for CHSRP and BSERP (*15, 16*)

Specification	CHSRP	BSERP
Origin	Sacramento–San Francisco	Beijing
Destination	Los Angeles–San Diego	Shanghai
Length (km)	1,280	1,318
Design speed (km/h)	350	350
Number of stops	25	21
Annual ridership	91–95 million passengers	160 million passengers
Construction cost	$45 billion (US)	$32.33 billion (US)
Unit cost ($million/km)	$35.16 (US)	$24.53 (US)

SOURCES: http://www.cahighspeedrail.ca.gov/; http://en.wikipedia.org/wiki/Beijing-Shanghai_Express_Railway.

In contrast, a number of high-ranking governmental officials serving on the Beijing–Shanghai Express Railway Leading Group oversee BSERP. The leading group consists of a vice premier of the State Council, cabinet-level ministers, and governors of the provinces along the HSR corridor and provides policy direction, oversees construction progress, and resolves major project issues involving multiple jurisdictions. Unlike CHSRP board members, who have equal status, each group member in BSERP has an unequal status, and the members form a vertical, hierarchical structure corresponding to each member's administrative rank in the government. In this group, the group leader (vice premier) often has final say on virtually all matters related to this project after the group meetings. Under this leading group, an office is housed in the MOR, which handles day-to-day routine business. The office chief is the minister of MOR. Therefore, the State Council leads and MOR manages BSERP, in conjunction with other governmental agencies and private firms. BSERP's planning and design were directly prepared by governmental agencies, especially MOR, subject to an independent review and certification by a third-party consultant, the China International Engineering Consulting Corporation.

Because of its small staff size, CHSRA contracted with private consultants and academic researchers to conduct various studies for the project (see Table 5 for details). CHSRA has the discretion to approve all planning documents, for example, environmental impact statements and environmental impact reports. In BSERP, the final authority to approve HSR plans and feasibility study reports rests with the State Council or goes through MOR, which reports to the National Development and Reform Commission.

To highlight the management features of HSR development, CHSRA is a state entity with coalition-building responsibilities, whereas BSERP followed a leading-group model, which is typical

TABLE 5 California High-Speed Rail Planning Activities

Date	Planning Activities
1997–2000	Detailed feasibility studies completed Corridor evaluation Ridership and revenue forecasts and cost/benefit analysis Project deliveries strategy Financial plan Business plan
2001–2005	Statewide program environmental impact report–environmental impact statement (EIR–EIS) studies completed Draft and final scoping reports Screening evaluation Environmental methodologies Alignment configuration and cross-sections Capital costs and operational and maintenance costs Economic growth effects Engineering criteria Operations Tunneling issues Statewide environmental technical reports for air quality, energy, and agriculture Regional environmental technical reports for biology, hydrology, land use, traffic, noise, parklands, visual impacts, cultural resources, geology, paleontological resources, public utilities, hazardous waste, and cumulative impacts
2004–2005	Implementation plan completed The implementation plan summarizes the California high-speed project—its alignment, stations, and technologies—as well as its financial and economic profile. It also lays out the roadmap for the CHSRA's evolution from a planning authority with a small staff to a construction management agency and, finally, to a comprehensive long-term manager of operations and assets.
2006–2008	Bay Area to Central Valley High-Speed Train Program EIR–EIS completed CHSRA and FRA prepared an additional program EIR–EIS

for the management of large projects in China. This difference in project management structure may be attributable in part to the fact that the jurisdiction of CHSRP falls within one state, whereas BSERP is a multijurisdictional project requiring the involvement of a higher than provincial-level government for project coordination and decision making. More importantly, this difference derives from the country-specific political system, which will be described in detail later in the paper. The Beijing–Shanghai Express Railway has the planning and design activities listed chronologically in Table 6.

Funding and Construction

Despite its strong political support and favorable ready-to-go project status, CHSRP faces a number of challenging issues. Most notably is the uncertainty over the funding from federal and private sources. The project faces a huge funding gap of more than $20 billion, according to a conservative estimate made by CHSRA. Table 7 shows the construction costs and funding sources for CHSRP and BSERP (*17, 18*). California Governor Arnold Schwarzenegger submitted California's application for federal stimulus dollars twice. During its Round 1 application, the funding requests included a total of $1.15 billion from ARRA Track 1, 3, and 4 grants for improvements to existing passenger rail lines that could eventually link up to CHSRP. In October 2009, the governor again submitted the funding requests for nearly $4.7 billion in federal ARRA funding for the engineering, design, and construction of CHSRP (Round 2 application). Track 1 projects can receive up to 100% federal funding with completed preliminary engineering and National Environmental Policy Act (NEPA) documentation. Other track projects require at least a 50% nonfederal match. On January 28, 2010, President Obama announced the provision of $8 billion for HSR projects across the United States, including $2.25 billion for CHSRP. This will leave a significant funding shortfall of between $25.8 billion and $28.3 billion for this project.

In contrast, BSERP is fully funded largely because of the lower project cost, strong backing from the central and provincial governments,

TABLE 6 Beijing–Shanghai Express Railway Planning and Design Activities

Date	Planning and Design Activities
December 1990	Conceptual alignment alternatives report of the Beijing–Shanghai Express Railway completed.
1994	Prefeasibility study of major techno-economic problems of the Beijing–Shanghai Express Railway completed.
December 1994	Feasibility study on BSER launched.
April 1996	Prefeasibility study report on the Beijing–Shanghai Express Railway (submitted for approval) completed.
April 1997	Supplemental pre-feasibility study report on the Beijing–Shanghai Express Railway (submitted for approval) completed.
October 1998–April 2000	Prefeasibility study report on the Beijing–Shanghai Express Railway (supplemental evaluation report) completed.
January 2000	Evaluation report on High-Speed Steel-Wheel-on-Steel-Rail versus Maglev Comparison completed.
2001	Notice on reserving land for the Beijing–Shanghai Express Railway issued.
July 2003–October 2003	International consulting for interim design guidelines completed.
September 2003	China International Engineering Consulting Corporation (CIECC) held public workshops on evaluating the necessity of launching BSER, modal alternatives, which concluded that steel-wheel-on-steel-rail alternative is the preferred alternative, not Maglev.
December 2003–July 2005	International consulting for design completed.
February 2, 2006	State Council approved the project proposal and confirmation. The project was officially incorporated into the NDRC's national plan.
May 2006–November 2006	CIECC completed the evaluation of the final project feasibility study report.
August 2007–October 2007	Project feasibility study report approved.
April 18, 2008	Construction began.

TABLE 7 Construction Costs and Funding Sources of CHSRP and BSERP (*16–18*)

Category	CHSRP ($)	BSERP ($)
Construction costs	45 billion	32.33 billion
Funding sources	State and local funding: 9.95 billion general obligation bond (Proposition 1A), approved by voters in state election on November 8, 2008 Federal funding: 2.25 billion from the ARRA high-speed rail award announced by President Obama on January 28, 2010 Public–private partnership (P3) funding: expect to raise 4.5 billion to 7 billion from project debt financing, vendor financing, system operations, and private ownership	MOR: 9.47 billion Local provinces, repayable insurance and pension funds: 7.36 billion Bank loans: 15.50 billion
Funding gap	25.8 billion to 28.3 billion	0.0

and a successful public–private partnership (P3). On December 28, 2007, MOR, along with provincial and municipal governments along the HSR corridor, inaugurated the Beijing–Shanghai Express Railway Company Limited, which functions as the project's construction arm. Strategic funding partners for the project include the China Railway Construction Investment Corporation, the investment and financing arm of MOR for the country's key railway projects; Ping An Asset Management Co., Limited, under the Ping An Group; the National Council for Social Security Fund; and the investment companies of the localities involved (*19*). The private sector has actively participated in project financing because of its high expectation and confidence that the HSR services between Beijing and Shanghai will likely generate enough ridership and revenues to repay the loans and costs.

At present, CHSRP has not entered the construction phase, pending environmental clearance, completion of the final design, and federal funding commitments. Should the project be kicked off before 2012, it will take about 10 years to complete Phase 1 construction, provided that major financial, environmental, and political hurdles are overcome in time. Otherwise, the project will be delayed. In the meantime, BSERP is already under construction. The Beijing–Shanghai Express Railway Company Limited, a shareholder-owned joint venture corporation, has selected four Chinese construction companies to undertake construction work after a competitive bidding process. The rolling stocks of BSERP will be manufactured by China CNR Corporation Limited and China South Locomotive & Rolling Stock Corporation Limited, with certain core technologies to be provided by foreign companies. Figure 3 shows a picture of the Beijing HSR Station, which has been in operation since late 2008 and which serves the Beijing–Tianjin HSR line. BSERP is scheduled to be completed by 2012.

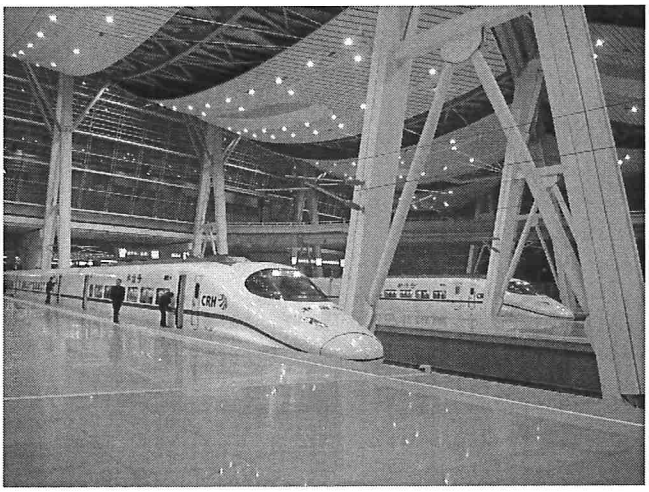


FIGURE 3 China Railway high-speed train.

INSTITUTIONAL ANALYSIS OF HSR DEVELOPMENT IN THE UNITED STATES AND CHINA

Political Systems

Politically, the United States is a federal constitutional republic characterized by a checks-and-balances governmental structure. The checks-and-balances governmental structure involves the separation of powers among the legislative, executive, and judicial branches and a vertical division of powers among federal, state, and local governments. The federal power has expanded dramatically over the 200 years since the United States was founded (*20*). Notwithstanding the dominance of the federal government, state governments do exercise a broad array of powers. Local governments in the United States have a large degree of autonomy and function to provide education, fire and police protection, street, land use control, sewerage, and sanitation services. They tend to be decentralized, fractionalized, fragmented, or polycentric (*21*). Largely reflective of the classic liberalism ideology, the U.S. political system is designed to prevent an individual or a governmental agency from possessing a monopolistic power, thus minimizing the potential for the abuse of power and corruption. This system may well result in the making of better and more prudent decisions, but it is often blamed for low efficiency, service duplication, and a waste of resources (*22*).

In contrast, China is a socialist republic in which the administrative responsibilities are largely held by a single party. The central government possesses a centralized power and often has the final say on almost all major policy decisions. Local governments form a hierarchical structure at the provincial, prefectural, county, and township levels. Whenever command and control from the top does not apply, fiscal institutions carry incentives to obtain the cooperation of the local governments (*23*). This political system often yields bipolar results. If the country has a visionary leader or a leading group with clear policy objectives, the system enables the efficient mobilization of the entire nation's resources to accomplish extraordinary projects. However, if the country falls into political turmoil, its citizens suffer in their personal and social lives and the national economy collapses.

Project Coordination Mechanisms

There is a tradition of planning and coordination across state lines in the United States. As required by the federal government, all urban

areas with populations of greater than 50,000 are subject to the planning, programming, and coordinating of federal highway and transit investments under the auspices of metropolitan planning organizations (MPOs). These MPOs coordinate the continuing, comprehensive, and cooperative planning process (*24*). However, these MPOs and regional planning remain fairly weak in the United States because of their lack of authority and resources, especially concerning land use, policy review, and transit operating policies. For these reasons, some scholars have labeled them "toothless tigers" (*25*).

In the rail industry, FRA is the key agency through which all HSR-related activities are funded and coordinated. FRA also serves as the federal lead agency for the HSR environmental review process for compliance with NEPA. However, as the subcabinet-level organization with an orientation toward the rail mode, FRA does not have enough authority in land use and other non-related-rail planning areas to overcome many organizational, multijurisdictional, and resource barriers (*26*). None of the state governors administratively reports to FRA. Under such a political structure, unless multistate or interregional agreements are reached, it is difficult, if not impossible, to coordinate planning activities and build consensus among various governmental agencies and stakeholders across multiple states.

In CHSRP, each board member has an equal status, which reflects the spirit of democracy and checks and balances. The coalition- and consensus-building processes, however, tend to be long because of the negotiation and persuasion activities required.

In China, major projects are usually managed by a leading group or commission formed specifically for the project (when the project is complete, the group or commission is dissolved). The leading group or commission coordinates interdepartmental and multijurisdictional activities. Depending on the importance and nature of the project, the position of group head or commission chair is held by governmental officials at the corresponding level. The BSERP leading group is chaired by a vice premier, indicating that the significance of the project is not at the same level as the Three-Gorges Dam project. Because provincial governors and ministers are lower in administrative rank than the vice premier, a vice premier-led BSERP leading group helps resolve interministerial, interdepartmental, and multijurisdictional conflicts, should any conflicts occur.

In summary, China has a centralized form of government, whereas the United States has a dispersed form of government, with the power in the United States being distributed among federal, state, and local governmental bodies. As reflected in project management, the Chinese leading group has an internal power hierarchy in which its members have unequal status and vertical working relationships, whereas the U.S. coalition-building group has an equal-status membership with horizontal working relationships.

Land and Railroad Ownership

In the United States, landowners have constitutional rights over their private properties. Acquisition of the rights-of-way needed for transportation projects has proved a time-consuming and financially costly process, even though the government can exercise its eminent domain power for the acquisition of land for public works. There are ample historical precedents in which local political resistance or litigation dreadfully delayed transportation projects that started with desirable initial public interest. For instance, a lawsuit against the Long Beach Freeway (Route 710) extension in Los Angeles County, California, by the city of South Pasadena has delayed the project for decades.

China has a public land tenure system: urban land is state owned, whereas rural land is collectively owned by villagers or farmers. This public land ownership offers legitimacy to the Chinese government at various levels to acquire land for public works projects. There have been an increasing number of protests from villagers and residents for being displaced by development projects. The protests mostly involve real estate developments. Local communities and villages in general welcome such projects as highways and rail lines or rail stations because they are expected to bring in resources and provide employment and service opportunities.

It is also worth noting that the higher-level governments in China have more control over local land uses than the U.S. federal and state governments. In the United States, local governments such as cities and counties are creatures of the state government and have the capacity of police power to control land use through zoning and subdivision regulations. The federal government has an indirect influence only over local land use decisions acquired through other related legislations, such as the 1969 NEPA, the 1970 Clean Air Act, the 1972 Clean Water Act, the 1991 Intermodal Surface Transportation Efficiency Act, and other more recent laws. In China, depending on the importance of a city, a local city's land use plan must be approved by the provincial government or the state council before the plan can be implemented.

In the United States, most railroads are privately owned and engage in freight train shipments. Amtrak is the sole intercity passenger railroad in the continental United States and operates passenger service on 21,000 mi (34,000 km) of tracks primarily owned by freight rail companies. This mixed system creates substantial difficulties for the planning, coordination, and implementation of new rail projects. For example, along some Northeast Corridor segments, the Acela Express train shares tracks with freight trains, significantly slowing down its operating speed.

In China, the government owns and operates both the railroad tracks and the rolling stock through MOR. This integration makes it relatively less painful to coordinate various rail-related activities and implement rail plans in an expeditious manner.

Environmental Impact Assessment Process

In the United States, any major investment projects must be assessed for their environmental impacts and cleared before they may proceed to the design and construction phases. Any federally funded projects, including HSR projects, must follow the rigorous NEPA guidelines for completion of the environmental impact assessment (EIA) process. For those projects funded by nonfederal sources, related state environmental laws must be followed as well. For example, CHSRP studies must follow the guidelines outlined in the California Environmental Quality Act, in addition to those outlined in NEPA.

In China, major investments also conduct EIAs using procedures similar to those used in the United States. However, China's EIAs tend to focus primarily on the pollution of air, water, and soil and give limited consideration to broader environmental, social, and health issues. Some local environmental protection agencies are funded by the development industries, which raises concerns over conflicts of interest. Consideration of project alternatives is at the minimal level. Many EIAs in China lack effective citizen participation. For example, the BSERP EIA process is primarily top–down, with little bottom–up feedback coming from ordinary citizens.

Funding Mechanism

P3 is essential to the success of HSR projects. P3 involves a contract between a public-sector authority and a private party in which the private party provides a public service or project and assumes substantial financial, technical, and operational risks for the project. Because the public sector is often unable to provide full funding for the project, the private sector may fill in the funding shortfall. The main rationale for public intervention in funding for HSR lies in the socioeconomic benefits linked to the HSR networks (*27*). Public funds can also attract private support because they reduce the investment risks for private investors (*28*). Public-sector and private investors have a synergistic relationship in P3 projects.

In the United States, because railroads are owned by profit-seeking private companies, many people believe that HSR projects should be self-financed by its riders without governmental subsidies, even though passenger transportation is a public business. In the past, three HSR corridor initiatives (Los Angeles–San Diego in California, Tampa–Orlando–Miami in Florida, and the Texas Triangle among Dallas–Fort Worth, Houston, and San Antonio) failed because the state and federal governments refused to participate and the projected net income was insufficient to support the enormous capital investment (*29*). To build a successful HSR network in the country, more governmental subsidies are urgently required. The federal government needs to develop a national transportation strategy and investment program to ensure that national funds are used to leverage funds from other governmental agencies and private investors (*30*).

In China, many private-sector companies, including pension funds, insurance companies, and investment companies, participate in funding BSERP as stockholders and thus share both risks and dividends. Hence, this type of P3 is an equity investment.

CONCLUSIONS

Amid an increasing globalization trend, both the United States and China are viewing the investment in and implementation of HSR as part of their national strategies to strengthen their economic competitiveness. In approximately 20 years, China has mobilized sufficient resources to turn HSR from a concept to a reality, with an additional 13,000 km of HSR tracks being laid down right now and in the near future. The United States has had decades of studies on HSR. It was not until the Obama Administration took office that federal funds were earmarked for HSR construction. Despite the political, economic, social, and cultural differences between the United States and China, the two countries can still learn from each other.

The high population density and relatively low personal income make HSR a desirable high-speed travel option for intercity trips for the majority of Chinese people. Rapid economic growth has generated large and growing travel demand among Chinese cities and regions. The urgent need for mobility improvement motivates the Chinese government to invest in HSR. In contrast, intercity travel in the United States is relatively well served by air and highway transportation systems, which creates a rather high threshold for HSR to enter the market.

China's centrally planned economic system enables the government to make major investment decisions and concentrate resources on HSR in a rather efficient way. Project-specific leading groups characterize the typical institutional strategy for the planning and implementation of large investment projects in China. The leading-group model is essential to resolving the interdepartmental and multijurisdictional conflicts that are common in large-scale infrastructure projects.

Even though China's institutional strategy for major project development cannot be directly transferred to the United States, the idea of having a strong leadership for HSR projects is still valid in the U.S. context. The federal government is obligated to lead the efforts for the development of the national strategic infrastructure. Stronger leadership backed by a clear vision helps boost the confidence of the private investors, who will then be more likely to enter into a P3. The Obama administration has demonstrated such a leadership role in its initial HSR plan. The resources that have been committed, however, are far from sufficient. In addition to empowering the federal government in the development of HSR, multistate partnerships and coalitions should also be strengthened to ensure that these agencies have actual teeth. To build public works projects like HSR, P3s are of the utmost importance. Otherwise, the acquisition of rights-of-way and tapping into private funding sources are extremely difficult.

The U.S. experience with HSR offers useful lessons to China as well. Most notably are the comprehensiveness of EIAs and the public involvement in the decision-making process. Ignoring the importance of these issues for the purpose of fast project delivery may lead to environmental and social backfire in the long run. Being fast is not equal to being good.

In a broader sense, the HSR developments in the United States and China are intrinsically related because major investments in HSR by the two countries will shift the demand curve in the international market for cement, steel, and technological know-how, which in turn affects the price and cost of HSR domestically. This is an important topic that warrants further studies.

REFERENCES

1. China Poised to Become World's High-Speed Rail Leader. *China Daily,* April 2, 2010. www.chinadaily.com.cn/bizchina/2009-09/10/content_8677419.htm.
2. Campos, J., and G. de Rus. Some Stylized Facts About High-Speed Rail: A Review of HSR Experiences Around the World. *Transport Policy,* Vol. 16, 2009, pp. 19–28.
3. Douchene, L. Probing Question: Why Don't We Have High Speed Rail in the US? April 27, 2006. www.physorg.com/news65373596.html.
4. *High-Speed Rail Strategic Plan: The American Recovery and Reinvestment Act.* FRA, U.S. Department of Transportation, 2009.
5. Wolman, D. Hey America, Make with the !@~$ High-Speed Rail Already. *Wired Magazine,* Issue 15.07, June 26, 2007. www.wired.com/cars/futuretransport/magazine/15-07/st_essay.
6. Thompson, L. S. High-Speed Rail (HSR) in the United States—Why Isn't There More? *Japan Railway and Transport Review,* No. 3, Oct. 1994, pp. 32–39.
7. Cox, W., and J. Vranich. *The California High Speed Rail Proposal: A Due Diligence Report.* Policy Study 370. The Reason Foundation, Los Angeles, Calif. www.reason.org/files/1b544eba6f1d5f9e8012a8c36676ea7e.pdf. Accessed July 15, 2009.
8. High-Speed Rail Corridor Descriptions. FRA, U.S. Department of Transportation. www.fra.dot.gov/us/content/203.
9. High-Speed Rail in China. *The Transport Politic.* www.thetransportpolitic.com/2009/01/12/high-speed-rail-in-china/.
10. High-Speed Rail in China. http://en.wikipedia.org/wiki/High-speed_rail_in_China.
11. California High-Speed Rail Authority. *Implementation Plan.* www.cahighspeedrail.ca.gov/images/chsr/20080123171537_ImplementationPlan.pdf. Accessed July 15, 2009.

12. California High-Speed Rail Authority. www.cahighspeedrail.ca.gov/images/chsr/20091026134234_Preferred_state_102209pm.pdf.
13. Beijing–Shanghai High-Speed Line, China. *Railway Technology.* www.railway-technology.com/projects/beijing/beijing1.html.
14. California High-Speed Rail Authority. Federal Stimulus Update. 2010. www.cahighspeedrail.ca.gov/library/default.aspx?name=Federal%20Stimulus%20Update.
15. Beijing–Shanghai High-Speed Railway. http://en.wikipedia.org/wiki/Beijing-Shanghai_Express_Railway.
16. California High-Speed Rail Authority. www.cahighspeedrail.ca.gov.
17. Beijing–Shanghai Express Rail Project Invites Bids. Economic and Commercial Section, Consulate General of the People's Republic of China in Gothenburg. http://gothenburg2.mofcom.gov.cn/aarticle/chinanews/200712/20071205272738.html.
18. Beijing–Shanghai High-Speed Railway. http://wapedia.mobi/en/Beijing-Shanghai_Express_Railway.
19. Beijing–Shanghai Express Railway Company Inaugurated. www.china.org.cn/english/news/237313.htm.
20. Whicker, M. L. The Functions and Structure of the Federal Bureaucracy. In *Public Sector Management* (M. L. Whicker and T. W. Areson, eds.), Praeger, New York, 1990.
21. Wikstrom, N. The Functions and Structures of State and Local Governments. In *Public Sector Management* (M. L. Whicker and T. W. Areson, eds.), Praeger, New York, 1990.
22. Stephens, G. R., and N. Wikstrom. *Metropolitan Government and Governance: Theoretical Perspectives, Empirical Analysis, and the Future.* Oxford University Press, New York, 2000.
23. Yang, J. Spatial Planning in Asia: Planning and Developing Megacities and Megaregions. In *Megaregions: Planning for Global Competitiveness* (C. L. Ross, ed.), Island Press, Washington, D.C., 2009.
24. Ross, C. L., J. Barringer, and A. A. Amekudzi. Mobility in the Megaregion. In *Megaregions: Planning for Global Competitiveness* (C. L. Ross, ed.), Island Press, Washington, D.C., 2009.
25. Feldman, P. Twenty-Five Years Later, SCAG Is Still a "Toothless Tiger." *Los Angeles Times,* April 8, 1991.
26. Meyer, M. D., and E. J. Miller. *Urban Transportation Planning—A Decision-Oriented Approach.* McGraw-Hill Higher Education, New York, 2001.
27. Thompson, C., and T. Bawden. What Are the Potential Economic Development Impacts of High-Speed Rail? *Economic Development Quarterly,* Vol. 6, No. 3, 1992, pp. 297–319.
28. Roll, M., and A. Verbeke. Financing of the Trans-European High-Speed Rail Networks: New Forms of Public–Private Partnerships. *European Management Journal,* Vol. 16, No. 6, 1998, pp. 706–713.
29. Southern California Association of Governments. *Southern California High Speed Rail Feasibility Study.* Southern California Association of Governments, Los Angeles, 1991.
30. Ankner, W. D., and M. D. Meyer. Investing in Megaregion Transportation Systems: Institutional Challenges and Opportunities. In *Megaregions: Planning for Global Competitiveness* (C. L. Ross, ed.), Island Press, Washington, D.C., 2009.

The Intercity Passenger Rail Committee peer-reviewed this paper.

Privatization Versus Public Works for High-Speed Rail Projects

Tsung-Chung Kao, Yung-Cheng (Rex) Lai, and Mei-Cheng Shih

With its virtues of high speed, large capacity, reduced levels of energy consumption, and low levels of pollution, high-speed rail (HSR) is emerging as an attractive transportation system. Because of the large investment burden required for HSR projects and the inefficiency of government-sponsored public works projects, many countries are now turning to the alternative of privatizing their HSR projects. The present research evaluates the success of privatized versus public HSR projects by comparing the outcomes for the Taiwan high-speed rail (THSR) project and the South Korean high-speed rail (KHSR) project. Except for the project delivery method (privatized versus public works), these two projects had similar project scopes and objectives and had parallel execution times. The results of the study indicate that a privatized HSR project such as THSR has a better likelihood of achieving traditional project management success in terms of time, cost, and quality; however, a government-sponsored HSR project such as KHSR could successfully promote the national HSR industry.

With the world's oil resources gradually becoming depleted and global warming developing into one of the worst environmental threats to the survival of humankind, transportation authorities are seeking remedies to the existing forms of energy used for transportation. Less energy consuming and more environmentally friendly green mobilization alternatives can replace the now heavily gasoline-dependent vehicles used for land and air transportation. Many countries are turning to high-speed rail (HSR) systems as a solution to meeting their transportation needs for the 21st century (*1*) (Table 1). However, because of the large investment required for construction of HSR projects, authorities have started to look into the alternative project delivery method of privatization. Furthermore, the use of privatization is also considered a means of improving the inefficiency often associated with a government-sponsored public works project (referred to as a government works project in this paper).

The Taiwan high-speed rail (THSR) project is the world's first privatized project of its kind. When the THSR project was initiated in 1989, the work was originally planned to be a government works project. However, after completion of the detailed design in 1993, the Taiwan government chose to undertake an unprecedented approach and privatize the THSR project. Privatization was a way to introduce the private owner's financial and management capabilities into a traditional public works project. Because of the use of this unusual approach, the THSR project was put on hold for a while to allow the government to complete the regulations required for a build–operate–transfer (BOT) scheme. After it was completed in 1998, the government awarded the HSR project to a private concessionaire, the Taiwan High Speed Rail Corporation (THSRC).

When Taiwan was under economic development pressure to construct HSR in 1990, South Korea also decided to build its HSR system to ease the congestion on highways due to its rapid economic development. Unlike the THSR project in Taiwan, however, the South Korean high-speed rail (KHSR) project was executed as a traditional government-sponsored public works project.

It is coincidental that the two projects had similar objectives and scopes and that both were parallel in their execution. The major difference was the method of delivery: a privatized endeavor versus a government works project. Therefore, comparison of the delivery methods of these two projects provides a unique opportunity to explore the effect of privatization on HSR projects.

THE PROJECTS

Taiwan High-Speed Rail Project

Project History

Taiwan's rapid economic development from the 1960s to 1980s greatly increased the demand on its intercity transportation system. As a result, the only freeway that served the island's north–south corridor became overly congested. This congestion not only reduced the level of service of the freeway but also prevented the further development of Taiwan's economy. Many business travelers turned to domestic airline services for their travel needs; however, these short flights (less than 30 min) are not environmentally friendly or efficient; overuse of the airports also hindered the planned cross-strait airline services.

Another important transportation provider on this corridor is the conventional railway, operated by Taiwan Railway Administration (TRA). Several alternatives were considered to improve rail services. Among those alternatives, upgrading of the conventional railway was carefully studied. To increase the train speed and shorten the travel time on the conventional railway, studies evaluated the possibility of widening the narrow-gage railway or introducing tilting trains. Both technologies, however, were considered either too expensive or too inefficient. The construction of a high-capacity HSR system to serve the country's transportation needs became the only viable alternative (*2*).

T.-C. Kao, Room 907; Y.-C. Lai, Room 313; and M.-C. Shih, Room 301, Railway Technology Research Center, Department of Civil Engineering, National Taiwan University, Civil Engineering Building, No. 1, Roosevelt Road, Section 4, Taipei 10617, Taiwan. Corresponding author: Y. C. Lai, yclai@ntu.edu.tw.

Transportation Research Record: Journal of the Transportation Research Board, No. 2159, Transportation Research Board of the National Academies, Washington, D.C., 2010, pp. 18–26.
DOI: 10.3141/2159-03

TABLE 1 Current Development of HSR

In Operation	In Construction	Planning
Japan	Netherlands	Poland
France	Iran	Portugal
Germany	Turkey	Russia
Italy		Morocco
Spain		India
Korea		Saudi Arabia
Taiwan		Argentina
United Kingdom		Brazil
Belgium		Indonesia
China		Vietnam
United States		

In 1989, the Taiwan government chose to proceed with the construction of the north–south HSR system and subsequently established the Preparation Office for THSR. After the completion of a detailed civil design in 1993, the government decided to privatize the project. This initiative aimed to allow the private sector to participate in the public works project so that it could bring an efficient management style and flexible financing capability to the project. In other words, the government hoped to eliminate the inefficiencies usually encountered in government works projects. In the long run, these inefficiencies were found to be responsible for project delays and the considerable escalation of costs. If the THSR project is successful in achieving the planned objectives, the Taiwan government would likely expand the privatization experience into other public works projects (*3*).

While awaiting the passage of legislation related to privatization laws, the construction of the THSR project was delayed for several years. In 1998, the THSR project was finally awarded to a private company, THSRC, which was to use a BOT scheme. The government and THSRC signed a construction and operation agreement (C&OA). This agreement to use a BOT scheme allowed the concessionaire (THSRC) to have the right to construct and operate THSR for a total concession period of 35 years. After the concession period, THSR will be transferred back to the government.

In 2000, a syndicated loan agreement of NT$312.4 billion (US$10.1 billion) between THSRC and a bank consortium was signed to fund the construction of the project. In the same year, the project's civil construction work commenced and the core system was awarded to the Taiwan Shinkanshan Corporation.

The commercial service for THSR was launched on January 5, 2007. After 2 years of service, THSR is now one of the mainstream transportation systems in Taiwan.

Project Scope

The scope of the THSR project undertaken by THSRC included the following (Figure 1):

- Build a 345-km HSR line from Taipei to Kaohsiung;
- Construct six new stations and retrofit two stations that THSRC shared with TRA;
- Establish five maintenance shops and depots at Sijhih, Lioujia, Wurih, Taibao, and Yanchao; and
- Purchase 30 train sets with maximum design and operating speeds of 315 and 300 km/h, respectively.

THSRC's Primary Goal

The 35-year concession period includes the construction and operation phases of the project. As a result, there was enormous pressure on the concessionaire to complete the HSR system and generate income from the fare box as soon as possible to pay for interest and depreciation. Therefore, the timely completion of the project to commence commercial service was the most important objective for THSRC.

Under the C&OA with the government, THSRC was authorized only to construct and operate THSR. In this case, there was no strong incentive for THSRC to acquire HSR technology through the construction of the THSR project. The only necessary technology transfer for THSRC was the technology for future operations and maintenance. Hence, the privatization of the THSR project led the concessionaire to focus on meeting its objective of achieving timely commercial service.

Implementation Strategy Affected by Privatization

The nature of this privatization project fundamentally affected the implementation strategy. The project's procurement strategy, organization, and management approaches were quite different from those of a public works project:

- On procurement strategy. THSRC used large lump sum design–build or engineering–procurement–construct contracts to delegate the design and construction responsibilities to contractors. Delegating responsibility to contractors streamlined the construction and design processes and reduced the interface responsibility required for THSRC. Consequently, the construction process was turned into one more like the purchase of shelved items.
- On project organization. THSRC turned to the international market to recruit temporary contract professionals.
- On project management. THSRC tried to delegate quality control to independent checking and supervising consultants.

Overall, all of these approaches aimed to accelerate construction and reduce THSRC's role in managing the construction.

South Korean High-Speed Rail Project

Project History

The rapid economic development of South Korea between the 1960s and the 1990s also caused congestion on the expressway between Seoul and Busan. Nearly 70% of the South Korean population and production capability were distributed along that corridor. Three options for improving traffic conditions were considered: a four-lane highway, a conventional double-track railway, and an HSR line. After careful evaluation, the HSR line was recognized to be the most efficient means of improving the situation. The decision to construct an HSR line linking Seoul and Busan was made in May 1989. In

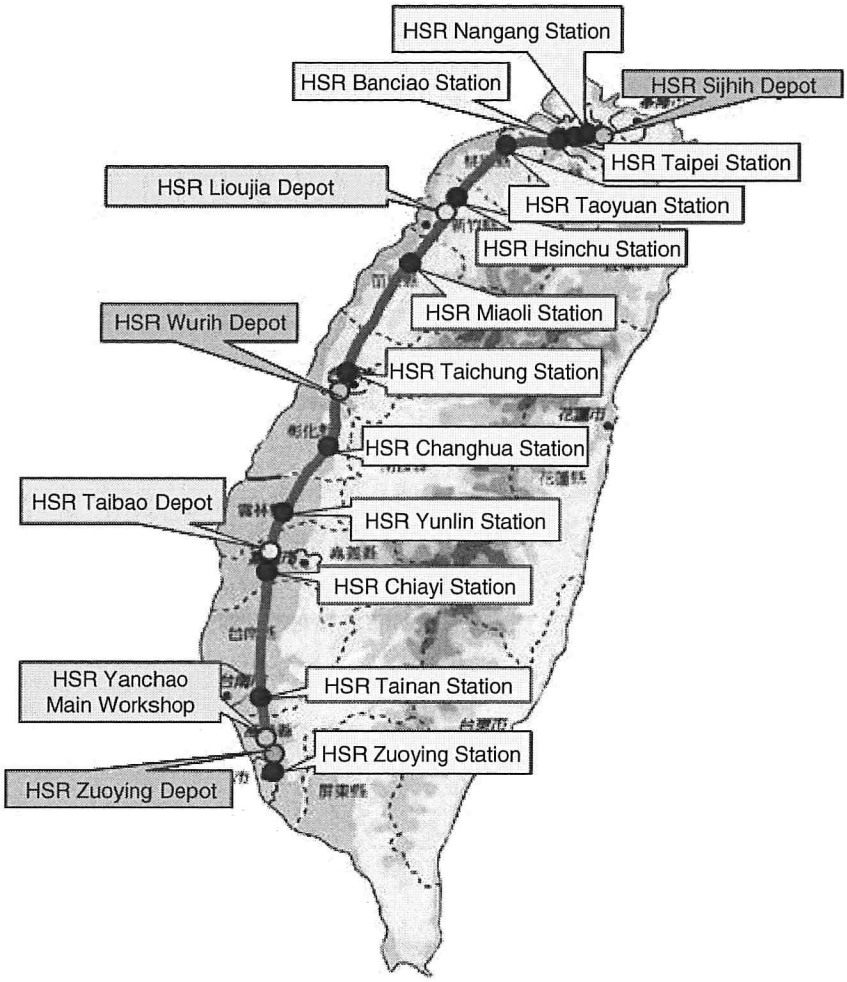


FIGURE 1 THSR network (2).

March 1992, the KHSR Construction Authority, which was to be in charge of the construction of KHSR, was established, and construction began on a 57.2-km test section in June of the same year (*4*). KHSR was carried out by the government as a traditional government-sponsored public works project.

During the construction of the KHSR project, numerous political and public interference and interruptions occurred; some of the most significant ones are summarized as follows (*5*):

- Route selection around Kyungju. Strong opposition to the proposed route through the outskirts of Kyungju mounted. Civic groups and environmental and cultural heritage experts, governmental organizations, and lawmakers took part in the debate, which lasted 3 years.
- Indecision over station location. Conflict erupted over whether the Daejeon and Dongdaegu Stations should be built underground or not next to existing railway stations. The construction plans for these two stations reversed repeatedly for nearly a decade.
- Naming a new station. Naming a KHSR station at the border of two cities (Choan and Asan) also became a long and difficult process.
- Tunnel under Mount Cheonsung. A Buddhist nun staged several hunger strikes in 2004 and 2005 to protest the construction of a tunnel under Mount Cheonsung. This incident consequently incurred large cost overruns and further construction delays.

Project Scope

The initial plan (Figure 2) was to construct an HSR line 412 km between Seoul and Busan. Four intermediate stations were located along the route. These four intermediate stations were Cheonan, Daejeon, Daegu, and Kyungju (*4*). The route included 112 km of at-grade sections, 109 km of viaducts, and 191 km of tunnels.

Because of the economic recession in the wake of the severe financial crisis that affected South Korea during the summer of 1997, the initial plan was revised in July 1998. Project implementation was divided into two phases:

Phase 1. Construct a new HSR line between Seoul and Daegu while implementing interim upgrades, and strengthen electrical connections to the existing conventional line between Daegu and Busan.

Phase 2. Construct the Daegu–Busan HSR line via Kyungju, as well as underground stations in Taejon and Daegu

The Phase 1 work was completed in April 2004, and the Phase 2 work was scheduled for completion in 2010.

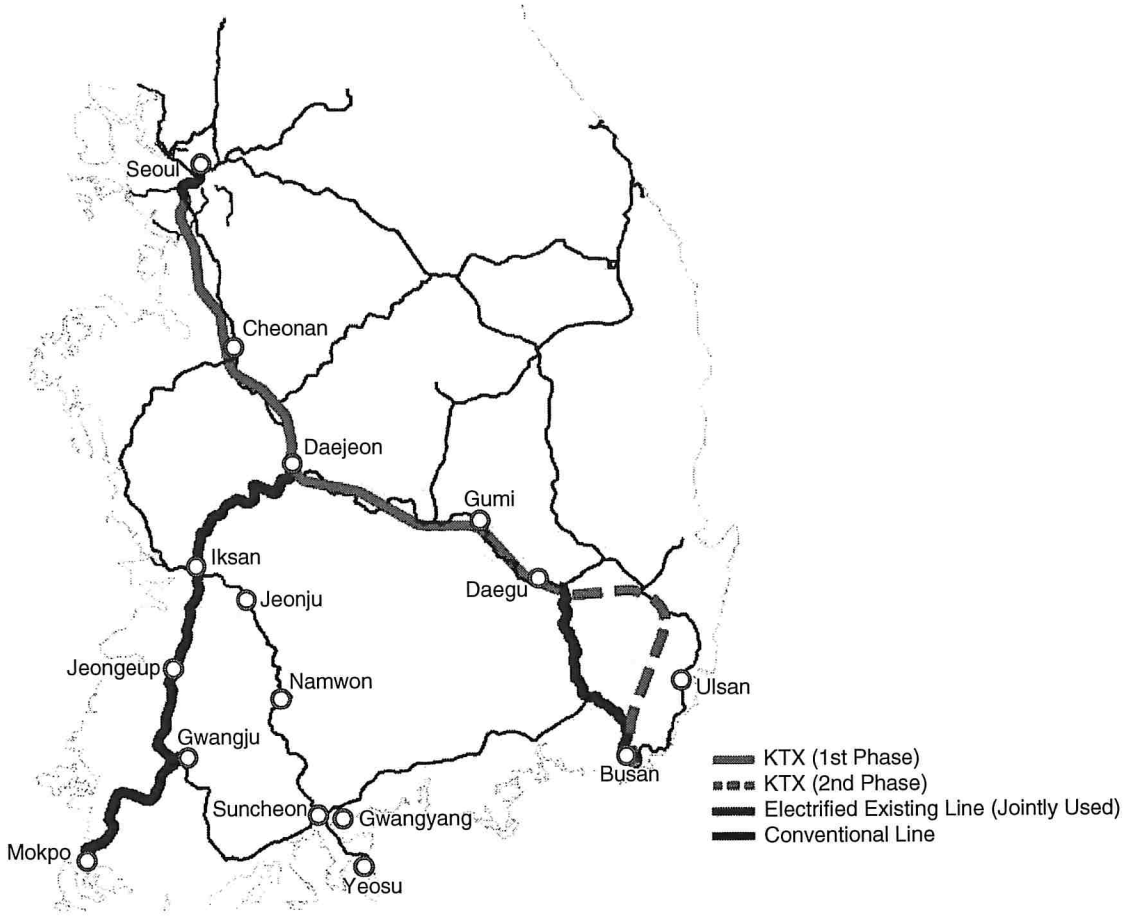


FIGURE 2 KHSR network (6).

Owner's (Government's) Primary Goals

The KHSR project had two primary goals (7). The first goal was to provide HSR service to the Seoul–Busan corridor. The second goal was to acquire the HSR technology.

Implementation Strategy Reflecting the Owner's Goals

To achieve the technical transfer of the high-speed trains, KHSR selected the French Train à Grande Vitesse technology in June 1994 and awarded the core system contract to a consortium led by Alstom. The scope of the contract included the supply of core systems and technology transfer. Among the 46 high-speed train sets, 12 were to be manufactured in France, whereas the remaining 34 train sets were required to be manufactured locally (4).

EVALUATION OF SUCCESSES OF PRIVATE AND PUBLIC HSR PROJECTS

A project is a temporary endeavor that takes place in a specific project environment to create a unique product. For practical reasons, projects cannot be repeated as laboratory tests to explore the effect of one distinct factor. Therefore, any comparison of different projects would be difficult because of differences in their objectives, environments, processes, and end products. However, the similarities that existed between the THSR project and the KHSR project provided a unique opportunity to explore the effects of the major differences in privatized versus government works projects. Both projects

- Were stimulated by a pressing economic need arising from Asia's rapid economic development between the 1960s and 1980s;
- Aimed at solving the freeway congestion that arose from the rapid economic growth;
- Planned to connect the capital city of each country with the largest ports in each country;
- Had similar route distances (345 km for Taiwan and 412 km for South Korea);
- Were aimed at achieving a top operational speed of 300 km/h;
- Were initiated in the same year (1990); and
- Were situated through the corridors with the most densely populated cities: THSR serves approximately 94% of Taiwan's population (22 million), and KHSR serves approximately 71% of South Korea's population (35 million).

It was essential to select the appropriate criteria to be used to evaluate the success of these projects. Over the past decade, many articles in the literature have proposed the criteria to be used to determine project success (8–12). On the basis of the existing work, the following

five criteria were chosen for use in the evaluation of the success of these HSR projects in this research:

- Time (the actual completion time versus the initial schedule),
- Cost (the actual cost versus the initial planned budget),
- Quality (system quality versus the initial planned quality),
- The owner's objectives (the satisfaction of the owners), and
- Users' needs (the satisfaction of the users).

The first three criteria (time, cost, and quality) define the success of project management (*8, 12*). Among these three criteria, quality is a rather vague term that requires a specific definition. For HSR systems, quality should be measured in terms of conformity to the functional and technical specifications. An HSR system integrates quite a few subsystems, including infrastructure, rolling stock, signaling and control, station operations, maintenance, management, marketing, and finance. Therefore, the quality of an HSR system depends not only on the quality of the subsystems but also on the integration among those subsystems. An approach for the investigation of quality would be based on the performance of the entire system under the designed operational requirements. Among all indices describing the performance of an HSR system, on-time performance is the most representative index because satisfactory on-time performance requires the excellent performance of all subsystems and their integration.

THSR has been in service for more than 2 years, whereas KHSR has been in service for more than 4 years. Both systems are now operating near their full capacities. As a result, the recent on-time performance criterion of each of these two systems can well represent the quality of their overall systems.

Past research has shown that the success of project management does not ensure the success of the product (*13*). Therefore, two more criteria, in addition to time, cost, and quality, were also considered: meeting the owner's objectives and satisfying users' needs (*13*). Meeting the owner's objectives can be evaluated by judging the completion of the owner's objectives in the initial project plan. In terms of satisfying users' needs, the increased rates in ridership for THSR and KHSR during the steady growth period were chosen as an indicator of users' satisfaction with the two new HSR systems.

When a transportation system is launched, users often require a period of familiarization until they start to adopt the new system; therefore, the buildup of ridership in the new system is often slow and is then followed by a rapid growth period. During the rapid growth period, the benefit and the convenience of the new transportation system gain widespread recognition by its new users. In addition, the new system may be integrated into the existing transportation network. After the growth period, the ridership gradually stabilizes because of market share stabilization with the shares for other competitive transportation systems. This ridership buildup phenomenon can be characterized by a logistic curve, as shown in Figure 3.

Time

During the construction of the THSR project, achieving the target date for commercial service was considered the most important objective of project management. THSRC was granted a 35-year concession period to construct and operate the HSR. In this case, any delay in launching the commercial service not only would result in additional construction costs but also would reduce revenue. The interest cost was initially estimated to be about US$3 million per day for the THSR project. Therefore, privatization of the THSR project compelled the project management team to maintain the timely goal of meeting the commercial service date as the most important target of the project.

THSRC signed the contract with the government to undertake the HSR project in 1998. Project construction began in March 2000, and the original plan was to launch the commercial service in October 2005; however, the commercial service was delayed until January 2007. The 14-month delay was mainly due to the late delivery of the core system.

The KHSR Construction Authority was created in March 1990, and construction began in June of the same year. The project was originally planned to start providing commercial service in December 1998. However, because of the delay stated earlier in this paper, the KHSR project was divided into two phases, with Phase 1 being completed and commencing commercial service in April 2004. Phase 2 is scheduled to be completed in 2010. The delay was about 63 months or more when only Phase 1 is considered.

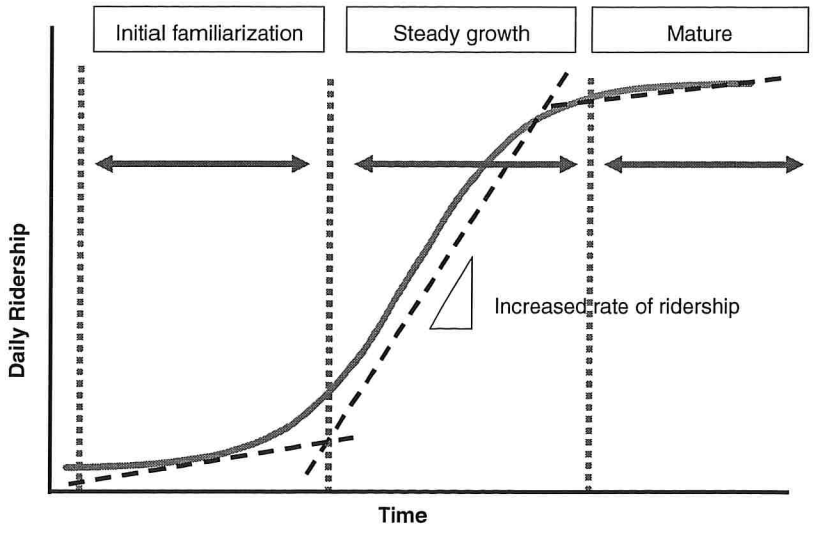


FIGURE 3 Typical trend in ridership increase.

Cost

The project cost overrun requires refinancing of the project by either raising the concessionaire's capital or obtaining additional loans from the financial institutions. Both of these measures could jeopardize the implementation of the project.

The public works projects carried out by the government are budgeted as an annual government expense; project cost overruns can often be allowed in the revision of the budget for the subsequent fiscal year. Public works projects have historically had more flexibility in accommodating project cost overruns. Therefore, cost control for many government-sponsored public works projects is less satisfactory than cost control for privatized projects.

The initial budget for the THSR project was estimated to be approximately NT$444.6 billion (US$14.4 billion) in 2000, and the estimated cost of the completed project was NT$472.7 billion (US$15.24 billion).

The budget for Phase 1 of the KHSR project was estimated to be approximately 5.85 trillion won (US$5.31 billion, in 2000 dollars) at the outset of the project, and the final completion cost was estimated to be approximately NT$12.74 trillion won (US$11.58 billion) upon commencement of system operation (*5*).

Quality

As stated earlier, the on-time performance of an HSR system is the indicator most representative of the quality of integration of all of its subsystems.

The on-time performance of THSR was reported to be nearly 99.4% after 2 years of commercial service (*2*). The on-time performance of KHSR was reported to be 95.5% for the same period of operation (*14*). The difference in on-time performance was partially caused by the fact that a portion of KHSR still shares tracks with the conventional railway.

Meeting the Project Owner's Strategic Organization Needs (Objectives)

The primary objective of the THSR project was to achieve commercial service with the forecast ridership of the system. For the KHSR project, the first objective was also to achieve the forecast ridership, and the second objective was to acquire the technology for an HSR system (*7*).

In this study, the average daily ridership of these two projects was compared with their forecast long-term goals to determine whether they achieved their first objective. For the THSR project, the forecast daily ridership was estimated to be nearly 300,000 passengers per day. For the KHSR project, it was projected to be approximately 150,000 passengers per day.

Figures 4*a* and 4*b* summarize the daily ridership on THSR and KHSR for the period from January 2007 to June 2009. Figure 5 compares the daily ridership on the two systems with their projected forecasts. After 20 months of operation, the THSR project achieved an average daily ridership of almost 84,000 passengers per day, and the KHSR project achieved an average daily ridership of 87,000 passengers per day. The rates of achievement of the long-term daily ridership were about 30% for THSR project and about 62% for the KHSR project. It was also interesting to discover that the average daily riderships of these two systems are similar to each other.

The second objective of the KHSR project was to acquire the technology required for the construction of an HSR system. This can be measured against the transfer of technology from Alstom to the KHSR project and can be evidenced by the high-speed train development capability of the South Korean industry.

Sharma reported that the technology transferred to the KHSR project involved the following (*15*):

- 350,000 documents,
- 23,000 pages in operation and maintenance manuals,
- 1,000 engineers being trained in France, and
- 400 French engineers working in South Korea.

Park (*16*) and Shin (*5*) reported that the South Korean industry has the following capabilities in the manufacture of HSR systems:

- A prototype high-speed train (KTX II) capable of running at a maximum speed of 350 km/h is being tested and will be used in the new HSR line (Seoul–Monkpo), which is under construction.
- A project to develop a high-speed train of the electric multiple-unit type capable of running at a maximum speed of 400 km/h to substitute for the current push–pull type high-speed trains was launched in July 2007 and is targeted to be completed in 2013.

On November 25, 2008, it was reported that KTX II was successfully launched in South Korea, and the KHSR project had achieved the ability to design and manufacture the high-speed train locally in South Korea. In this regard, one can say that the KTX project has achieved the objective of acquiring the HSR technology.

Table 2 summarizes the performance of the two projects in meeting the owners' objectives.

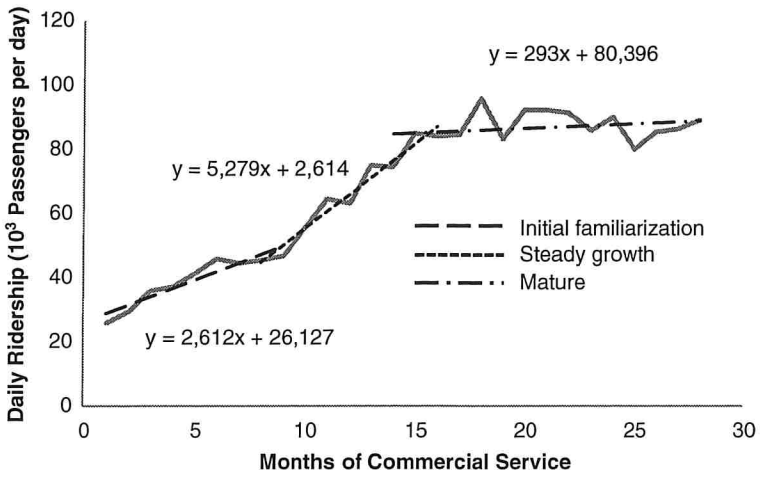
Satisfying Users' Needs

As stated earlier, the increased rate of the ridership during the period of steady growth of a transportation system could be used as an indicator to measure the satisfaction of the users' needs for that system. As shown in Figure 4, for the THSR project there was an increased rate of daily ridership of approximately 2,600 passengers per day per month in the initial familiarization period, and this was followed by a rate of about 5,300 passengers per day per month in the steady growth period.

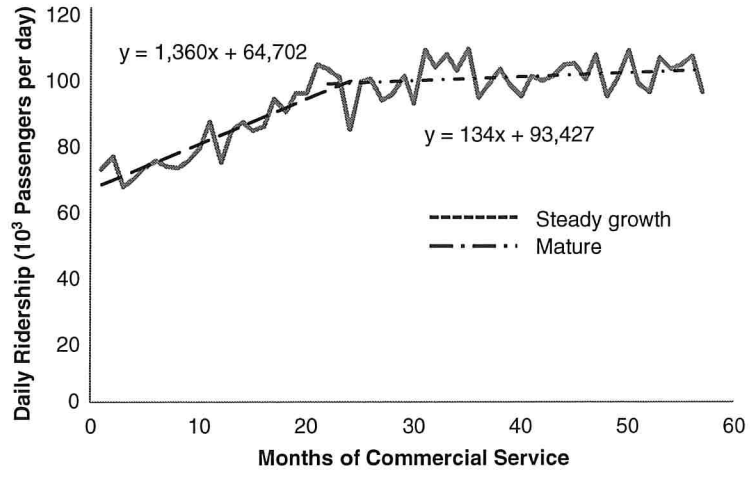
On the other hand, on KHSR the increase in ridership to its mature phase was very rapid. As indicated in Figure 4, the growth curve for KHSR ridership shows only the steady growth and mature periods; there is no evidence of an initial familiarization period. According to Figure 4, the increase in the rate of ridership on KHSR was approximately 9,400 passengers per day per month in the steady growth period; this was followed by about 1,330 passengers per day per month during the mature period.

COMPARISON AND LESSONS LEARNED

In reviewing the criteria used to determine the success of project management for these two projects, it was evident that the THSR project performed relatively better in terms of the traditional project objectives of time, cost, and quality. However, the government-sponsored public works project of KHSR achieved better product success in terms of meeting the owner's strategic needs (objectives)



(a)



(b)

FIGURE 4 Reported ridership of (*a*) THSR and (*b*) KHSR.

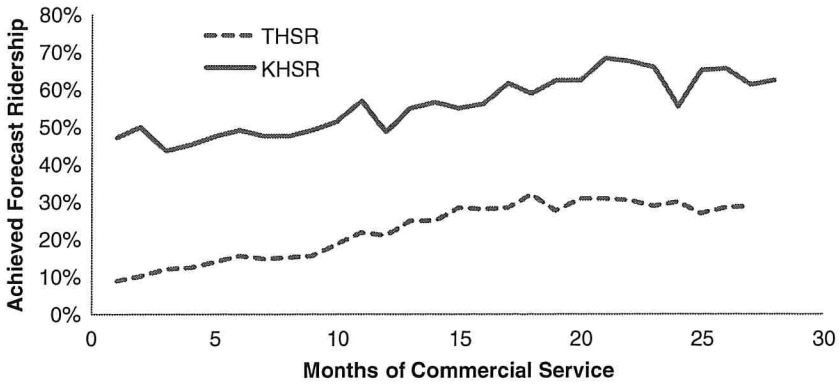


FIGURE 5 Percentages of ridership goal achievement for THSR and KHSR.

TABLE 2 Comparison of THSR and KHSR Owners' Strategic Needs

	Strategic Needs	
HSRs	Objectives	Portion of Objectives Achieved
THSR (Phase 1)	The commercial service	About 28% of forecast ridership
KHSR	Provide additional rail service in Seoul–Busan corridor	About 60% of forecast ridership
	Acquire technology for high-speed system	Achieved successfully (100%)

and satisfying the users' needs. Detailed comparisons and discussions of these two components of the success of these two HSR projects are presented below.

Time

The first time that both Taiwan and South Korea undertook the construction of an HSR system occurred during the projects described here. The technical challenges for both projects were quite similar. However, the completion time for the THSR project was better than that for the KHSR project.

In reviewing the construction history of the KHSR project, it was evident that the construction of the project was heavily affected by the external interference caused by political rivalry and public objections. External interference included the indecision over the alignment of a section of the route for 3 years, the flip-flop decisions over station locations, and the massive public objections to the tunnel under Mount Cheonsung inflamed by a single individual.

In the past, these kinds of influences also caused serious delays to government works projects in Taiwan. At times, this interference was motivated by genuine concerns over public interests. More often, however, they were mediated by hidden personal agendas. One of the reasons that the Taiwan government decided to privatize the THSR project in 1993 was to avoid the effects of these unnecessary influences.

During the construction of THSR, interference similar to that described above occurred every so often. Political pressures were mounted on the governmental agency responsible for the THSR project and on THSRC to realign the route, to relocate the stations, and to suspend construction. Because the THSR project was privatized, the Bureau of High-Speed Rail (the governmental agency responsible for the project) was able to avoid these political pressures because it was not in charge of project implementation. In a few incidents, THSRC was also able to call on the governmental executive power to resist the public objections aimed at interrupting the construction of the project. It is evident that privatization of the THSR project successfully curtailed the external influences that interfered with the project.

In government-sponsored public works projects, the objective of meeting the project's original time schedule was often predisposed to political rivalry and personal interests. The ability to manage and safely guard a project's time schedule was compromised. Accountability for the project's time schedule was also ambiguous. Privatization of the THSR project had ultimately allocated the accountability for achieving project time control within the hands of the concessionaire.

On the basis of the ability of these two projects to control their project times, it can be concluded that a privatized HSR project provides a better environment for management of the time schedule.

Cost

Similar to the inability to control the project schedule, the project management team of a government-sponsored public works project could also face situations over the project cost beyond its control. Consequently, the cost of a government works project was inflated by project delays and changes caused by external interference.

In its eagerness to gain support from legislative institutions and the public, the time and cost overruns of the government-sponsored public works project were further aggravated by the project team's underestimation of the time and costs at the project's outset. Therefore, underestimates were often made with an understanding of the ambiguity of the accountability as part of the nature of government work.

Revision of the budget of a privatized project, however, requires the approval of the investors in the project and the lending financial institutions. These approvals were more difficult to achieve than revision of the government's fiscal budget. Therefore, because of the stringent project control requirements of a privatized project, one can also expect that a privatized HSR project can better manage project costs.

Quality

The on-time performances of both projects were impressive (both were over 95%). Because the two systems were able to operate with satisfactory on-time performance during their operation at the required maximum operating speed (300 km/h) and at their design capacities, it can be concluded that privatization does not have an impact on the quality of HSR projects.

Meeting the Owners' Strategic Goals

For a privatized HSR project such as the THSR project, the primary goal for THSRC was to generate sufficient income from ridership to pay for the investment. However, for the government-sponsored KHSR project, an additional goal was the South Korean government's expectation to develop the HSR industry.

Comparison of the owners' strategic goals for ridership achieved after 2 years of operation showed that the KHSR project achieved its forecast long-term ridership at a higher level than its counterpart, the THSR project (60% and 28%, respectively). However, comparison of the average daily ridership figures for these two HSR systems after 2 years of operation showed that the two were close in the average number of passengers (87,000 and 84,000 passengers per day, respectively). Nonetheless, the expected ridership for THSR was twice as high as that of KHSR (300,000 and 150,000 passengers per day, respectively). The significant difference in ridership expectations

may arise from the need for the owner of the privatized project to impress its investors and lending financial institutions.

With the experience of constructing its first HSR and the transfer of technology from its French partner, South Korea can definitely be proud of becoming a country with HSR and a manufacturer of high-speed train systems. In the beginning, the Taiwan government also had the ambition of developing its rail industry through the implementation of the HSR project, but the privatization of the project provided no incentive for the concessionaire to pursue this initiative.

Satisfying Users' Needs

As stated earlier, the increased rate of ridership during the period of steady growth of a new transportation system can be used as an indicator of the users' satisfaction with the new system. Comparison of the two new HSR systems indicated that the growth in THSR's ridership was more deliberate than that of KHSR. Furthermore, the comparison also showed no evidence of an initial familiarization period for the KHSR project.

The difference between the ridership growth patterns of these two HSRs can also be attributed to the privatization of the THSR project. The initial ridership of the THSR project was much lower than that of the KHSR project (30,000 and 70,000 passengers per day, respectively). Surveys conducted by THSRC at the outset of its operation to investigate the market reaction indicated that the inconvenience of transfer connections was one of the major reasons for the slow pace (*2*). The inconvenience was partially caused by the need for THSR to gain cooperation from other public transportation systems so that passengers could complete their full journeys. However, these public transportation systems were also THSR's competitors, and the introduction of THSR was somewhat viewed as a disturbance to the status quo.

On the other hand, the KHSR system is part of the South Korean Railroad (Korail). The intermodal connections for passengers were already well established by the existing Korail system. There was no resistance to KHSR's introduction into the existing transportation network. As a result, KHSR's commercial service can be viewed as an extended service. On the other hand, a brand new privatized HSR system such as THSR may require a longer period of time to build up its ridership.

CONCLUSION

The privatization of the THSR project presented a unique opportunity to explore the effects of privatization on HSR projects. The study showed that privatization of the THSR project was able to curtail the external influences that might impede the implementation of a project. Moreover, a privatized HSR project has a better chance of achieving the traditional project management success according to the criteria of completion time, cost, and quality.

However, the privatized THSR project indicated that the ridership suffered a slower rate of growth during the initial period of operation because of unfamiliarity and a lack of convenient connecting transportation systems. This type of HSR project will therefore require a prolonged period to be integrated into the existing transportation network.

The government-sponsored KHSR project, however, successfully promoted the national HSR industry and achieved a wider scope of product success than the THSR project. Consequently, a privatized HSR project can be considered a commercial investment for the concessionaire, whereas a government-sponsored HSR project may be seen as a national campaign.

REFERENCES

1. *Vision for High-Speed Rail in America.* FRA, U.S. Department of Transportation, 2009.
2. Taiwan High Speed Rail Corporation. *Company Overview.* http://www.thsrc.com.tw./en/about/ab_comp.asp. Accessed July 1, 2009.
3. Research, Development, and Evaluation Commission. *Evaluation of the BOT Project in Taiwan.* Executive Yuan, Taiwan, 2006.
4. Kang, K. D. Construction of the Korean High Speed Railway. *Rail International,* 2000, pp. 19–23.
5. Shin, D. C. *Recent Experience of and Prospects for High-Speed Rail in Korea: Implications of a Transport System and Regional Development from a Global Perspective.* Institute of Urban and Regional Development, University of California, Berkeley, 2005.
6. Lee, J. H., and J. S. Chang. The Impact of High-Speed Rail Service on the Modal Share of Intercity Passenger Travels in Korea. Presented at 85th Annual Meeting of the Transportation Research Board, Washington, D.C., 2006.
7. Suh, S. D., and K. Y. Yang. Customers' Reactions to the Introduction of High-Speed Rail Service: Korean Train Express. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1916,* Transportation Research Board of the National Academies, Washington, D.C., 2005, pp. 20–25.
8. Might, R. J., and W. A. Fisher. The Role of Structural Factors in Determining Project Management Success. *IEEE Transportation in Engineering Management,* Vol. EM-32, No. 2, 1985, pp. 71–77.
9. McCoy, F. A. Measuring Success: Establishing and Maintaining a Baseline. *Proc., PMI Annual Seminar & Symposium,* Montreal, Quebec, Canada, 1986.
10. Wateridge, J. How Can IT/IS Project Be Measured for Success? *International Journal of Project Management,* Vol. 16, No. 4, 1985, pp. 59–63.
11. Wells, W. G. From the Editor. *Project Management Journal,* Vol. 29, No. 4, 1998, pp. 4–6.
12. Baccarini, D. The Logical Framework Method for Defining Project Success. *Project Management Journal,* Vol. 30, No. 4, 1999, pp. 25–32.
13. Collins, A., and D. Baccarini. Project Success—A Survey. *Journal of Construction Research,* Vol. 5, No. 2, 2004, pp. 211–231.
14. Korail. info.korail.com/2007/eng/eng_index.jsp. Accessed July 1, 2009.
15. Sharma, V. Welcome Address. *Proc., 5th Training on High Speed Systems,* International Union of Railways, 2008.
16. Park, J. High Speed Rail System in Korea. *Proc., 5th Training on High Speed Systems,* International Union of Railways, 2008.

The Intercity Passenger Rail Committee peer-reviewed this paper.

Energy Consumption and Emissions of High-Speed Trains

Alberto García Álvarez

This paper analyzes the relationships between the speed of passenger trains (especially high-speed trains), their energy consumption, and their greenhouse gas emissions. After an outline of the special energy consumption characteristics of high-speed railway systems, the amounts of energy consumed by conventional and high-speed passenger railway systems are compared. The comparison includes an empirical verification of the differences between high-speed and conventional rail systems and an analysis based on theoretical models. It is shown that, on average, high-speed railway systems usually consume 29% less energy than conventional railway systems. With a comparison of the levels of energy consumption and emissions of high-speed passenger trains with those of all other modes of transportation with which it competes (including conventional passenger trains), the net effects on emissions of high-speed train service on any corridor in the study can be analyzed. This is important because even if the difference in the energy consumption of the Spanish high-speed rail system, Alta Velocidad Española (AVE), and that of conventional rail system is not significant or even if AVE consumes more energy, the diversion of passengers from air travel ultimately yields significant reductions in energy consumption and emissions on a route. The study concludes that each high-speed train passenger accounts for an emissions reduction of approximately 30 kg of CO_2 and that this reduction increased on the routes on which AVE reaches higher speeds.

The emergence of a high-speed passenger rail system in Spain has allowed railways to significantly increase their market share of passengers traveling over medium and long distances (75 to 700 km). This increase in the number of passengers has been achieved by attracting both travelers away from other modes of transportation and new users because of the increased mobility that it offers. Moreover, expansion of the high-speed rail network and high-speed rail services has given rise to new patterns in urban development and improvements in the quality of life of many people.

However, high-speed rail is also perceived to have a number of disadvantages, such as the need for large investments, the supposedly high operating costs, and an unacceptably high level of energy consumption (and, therefore, high levels of emissions).

Some studies, such as the study of Minayo and García (*1*), have shown that non-energy-related operating costs (in particular, the costs for rolling stock, maintenance, and personnel) not only do not increase but also actually decrease with increasing speeds.

Fundación de los Ferrocarriles Españoles, Santa Isabel 44, Madrid 28012, Spain. albertogarcia@ffe.es.

Transportation Research Record: Journal of the Transportation Research Board, No. 2159, Transportation Research Board of the National Academies, Washington, D.C., 2010, pp. 27–35.
DOI: 10.3141/2159-04

PURPOSE OF THIS PAPER

This paper analyzes the energy consumption of high-speed railway systems and seeks to answer the following questions:

- Does the high-speed railway system entail an increase in energy consumption, per passenger, in comparison with that of the conventional railway system?
- Does the high-speed railway system require less energy, per passenger transported, than other modes of transportation? What is the order of magnitude of the difference?
- Does the development of a high-speed railway on a route cause a net increase or a net decrease in the amount of energy consumed for transportation?

In short, the objective is to analyze whether expansion of the high-speed rail network is going to exacerbate the problem of carbon emissions and energy consumption in the transportation sector or whether transportation by high-speed rail is more energy efficient than transportation by other, alternative modes (including conventional rail) and, therefore, whether investment in high-speed rail lines and trains will help improve the sustainability of the transportation system.

GENERAL COMMENTS

This subject has no absolute or universally valid truths. This paper analyzes specific cases and corridors that correspond to the Spanish sphere, from which 16 years of experience with high-speed rail operations provides conclusive data. In another type of operation, in other geographic environments, or at other moments in time, the results might be different.

The author has published various articles on the subject, with special emphasis on comparison of conventional and high-speed trains (*2, 3*) and comparison of high-speed rail with other modes of transportation (*4*). This paper summarizes, completes, and updates the results of that work and, at the same time, adds new data and further reflections on the impact of high-speed rail on the levels of energy consumption by the transportation system.

COMMON MISCONCEPTIONS ABOUT ENERGY CONSUMPTION BY HIGH-SPEED RAIL

There is undoubtedly a widespread belief among the general public that the operation of a high-speed train requires a lot of energy. This common misconception serves as a justification to attack high-speed railways and is often used by the opponents of high-speed rail as a reason against constructing new high-speed rail lines, for improving

conventional lines, or simply, for strengthening other modes of transportation.

An example of this belief can be found in some of the entries that Spanish schoolchildren submitted to a competition held by the Railway Museum in 2006. To find the information that they needed, the youngsters searched the Internet and found the ideas that they believed were the most suited to the task.

Several entries coincided, and an example is taken directly from an entry submitted by fourth-year students at a secondary school in Madrid: "The high-speed train consumes a disproportionate and unsustainable amount of energy. AVE [Alta Velocidad Española, the Spanish high-speed rail system] that runs between Madrid and Seville at a maximum speed of 300 km/h (at an average of 209 km/h) has a power output of 8,000 kW: it consumes as much electricity, measured in kilowatts per hour, as a city of 25,000 inhabitants."

Two ideas relate to the links between energy, power, and the operation of trains that are widespread, even among many engineers and railway experts:

1. The "power rule" or the "cube rule" is the idea that the power that the train needs to overcome aerodynamic drag increases with the cube of its speed.
2. The "energy rule" or the "square rule" indicates that the level of energy consumption increases in proportion to the square of the speed because the amount of energy consumed is proportional to the aerodynamic drag.

LACK OF INFORMATION ABOUT ENERGY CONSUMPTION OF HIGH-SPEED TRAINS

Roger Kemp, a professor of engineering at the University of Lancaster (Lancaster, Great Britain), has repeatedly insisted on the energy-related harmfulness of high-speed trains with affirmations such as the one contained in his provocatively titled article "Take the Car and Save the Planet" (*5*), in which he obtains various comparative results that do not favor the use of passenger high-speed trains at all.

The quantitative basis of that professor's claims is the consumption attributed to a hypothetical high-speed train running at different speeds between London and Edinburgh in the United Kingdom (*6*). With conditions of 57 kWh of energy used per seat for the whole journey, a joint train and transmission efficiency of 65%, and a power station efficiency of 0.40, Kemp calculates that the amount of fuel consumption per seat is 22 L on the train journey from London to Edinburgh when the train is traveling at a speed of 350 km/h. He then compares the fuel consumption of the high-speed train with that of a Volkswagen Passat 130 TDI automobile carrying two occupants, to which he attributes the amount of fuel consumed to be 2.8 L/100 km (and a well-to-tank efficiency of 80%) and reaches the conclusion that the levels of consumption are the same (22 L per passenger) by the two modes of transportation. His comparison is completed by analysis of the fuel consumption of a single-class Airbus 321-100 aircraft, to which he attributes the amount of fuel consumed to be 16 L per seat and for which he uses the same well-to-tank efficiency as that for the automobile (80%), resulting in the consumption of 20 L per seat for the journey from London to Edinburgh by air. In other words, the amount of fuel consumed by the aircraft is less than that consumed by the high-speed train.

It would seem advisable to point out that Kemp has not taken into account the fact that the same amount of primary energy used to produce electricity does not have the same effect on greenhouse gas emissions or on the contribution of fossil fuels to the exhaust because the primary energy source for transportation is oil. Therefore, the mere comparison of primary energy is not relevant for any practical purpose.

Nevertheless (and disregarding numerous methodological aspects that could be called into question), the most important difference between the studies of the author of this paper and the study of Kemp lies in the amount of fuel consumed by the high-speed train, as Kemp's estimate does not seem to tally with the actual amounts consumed by high-speed train operations in Spain over the last 16 years (*6*).

Consider the route from Madrid to Barcelona, Spain (620 km, which is somewhat longer than the London–Edinburgh route), and take a Series 103 high-speed train with 397 seats. The journey time was calculated with an Aplica simulator, and the fuel consumption was calculated with Alpi2810 software (both of which were extensively checked against the actual figures for Spain). To ensure homogeneity, the time calculation is based on a margin in the middle of the International Union of Railways band (8%), and fuel consumption is calculated on the basis of the energy sent to the train's wheel rims (including the auxiliary services) and for each train seat, so that the method used to calculate the data is strictly comparable to that used previously (*6*).

Despite the greater distance (+3.3%), the results show that the journey time calculated for Madrid to Barcelona is slightly shorter than that calculated by Kemp (between 6 and 16 min), which can be attributed to logical differences, such as the layout of the line or the margins adopted. The results for the journey from Madrid to Barcelona also show that the amount of energy consumed is considerably smaller than the amount calculated by Kemp, with the differences being up to 56%. The Series 103 train from Madrid to Barcelona running at 350 km/h consumes 25.08 kWh per seat, on the basis of calculations performed with the energy sent to the wheel rims, less than half the amount presented by Kemp, although it should be pointed out that the existence (as in reality) of a regenerative brake was assumed, whereas this advantage would almost certainly not have been considered in Kemp's study (which was based on a 1981 French Train à Grande Vitesse, which did not have a regenerative brake). Yet even without consideration of the contribution of the regenerative brake, the amount of energy consumed by AVE traveling at 350 km/h between Madrid and Barcelona would be 27.47 kWh per seat, 52% less than the amount indicated by Kemp (*6*). All amounts refer to the numbers of kilowatts per seat on the complete 620-km route.

COMPARISON OF ENERGY CONSUMPTION OF HIGH-SPEED AND CONVENTIONAL RAILWAY SYSTEMS

In view of the background presented above, an attempt will be made to compare the energy consumption and greenhouse gas emissions of a high-speed passenger railway system with those of a conventional passenger railway system. This comparison is based on empirical verification and the results of analyses carried out with simulators, with the aim being to explain the reasons for the various differences in levels of consumption that have been calculated.

The comparisons entail a difficulty deriving from the fact that the cases being compared are differentiated by other variables (besides speed). In fact, the high-speed trains and the lines on which they run are different from the conventional ones; and the differences, besides speed, relate to the layout, the power supply, the number of stops, and so forth.

High-speed rail systems require certain curve-free layouts, lighter and more aerodynamic trains, and a power supply system that can provide high outputs. Without these attributes, high speeds would not be possible, and therefore, the comparison must incorporate all these attributes. In short, the comparison of the levels of energy consumption by the various trains is made by use of the conditions of the normal operating environment of each one: the characteristic lines, their sizes, their occupancy rates, and so forth.

Empirical Verification

An initial approach involves observing reality from a general perspective to find out whether the common misconceptions have any basis in reality or if, on the contrary, reality offers different results.

Empirical Relationship Between Power Output and Speed

According to the common misconception relating to the train's power output, the output would increase with the cube of the train's speed.

To find out whether this is actually true, the outputs of various electric traction units of Spanish railways at their maximum speeds were compared. In the case of the electric traction units, those with an output capable of pulling trains consisting of six or seven passenger cars were considered; that is, the trains had a transportation capacity equivalent to that of the modern high-speed trains with which they are being compared.

In no case does the cube rule apply: the outputs are always much lower than they should be according to the cube rule. If high-speed trains traveling at 350 km/h complied with the cube rule in relation to the electric traction units that were used in 1923, the rate would be $350/90^3$, which would be 59 times greater; therefore, the high-speed trains should have a power output of nearly 70 MW, whereas they actually have an output of 8.8 MW.

The relationship between each vehicle's output (in kilowatts) and its maximum speed (in kilometers per hour) shows a curve that is certainly not cubic and whose equation is

$$P = 4{,}468 \times S^{1.3021}$$

where P is power (kW) and S is speed (km/h).

Empirical Relationship Between Energy Consumption and Speed

Regarding the common misconception about energy consumption (i.e., the square rule, in which the energy consumption of trains increases with the square of their speed), various types of trains that cover a variety of real situations were compared.

An initial comparison can be obtained from the Renfe 2005 report (*7*) by extracting the data corresponding to traction energy use and traffic data for each one of the business units (commuter, regional, long-distance, and high-speed trains). The average speeds are relatively different from one unit to another, but there is a certain degree of uniformity within each unit.

The high-speed business unit, the one whose trains run at the highest speed (160 km/h), is precisely the one that records the lowest amount of energy consumption per passenger kilometer [3.39 W-h per passenger kilometer (W-h/p-km)], whereas for the rest of the business units, the amount of energy consumed per passenger kilometer increases as the average speed decreases. The scores for the other units are as follows:

- Commuter rail: 53.29 km/h and 130.3 W-h/p-km,
- High speed rail: 159.99 km/h and 73.39 W-h/p-km,
- Long distance (no high-speed rail): 89.29 km/h and 82.25 W-h/p-km, and
- Regional rail: 71.21 km/h and 93.04 W-h/p-km.

Articles Based on Specific Cases

In recent years, the author has published various articles about the energy consumption of all types of high-speed trains. Some of these articles, together with others produced by other authors, allow the energy consumption of the high-speed train to be compared with that of the conventional train in specific cases, which are described in the following sections.

Different Trains on Different Lines

The aim of the first of these articles was to analyze the energy guzzler accusations that have been leveled against high-speed trains (*2*). The energy consumption of a conventional train running at 200 km/h on an improved conventional line (Barcelona to Alicante, Spain) was therefore compared with the energy consumption of a high-speed train on the Madrid–Lleida, Spain, high-speed line. In both cases, the trains have 316 seats.

The average speed reached by the high-speed train is 32.6% higher, even though it consumes less energy per kilometer, both when the energy imported to the pantograph is measured (–7.2%) and when the net energy (when the energy exported to the network is deducted) as it leaves the power plant is recorded (–15.7%). Given the objective of the experiment, the effect of the shorter length of the high-speed lines in comparison with the effect of the length of conventional lines was not taken into account, which would make the difference even greater.

Despite a lower average speed, the level of energy consumption is higher in the improved conventional train both in the train's wheel rims and at the power station outlet. Certainly, the level of consumption due to aerodynamic drag is much higher (+71.6%) for the high-speed train, but the levels of consumption for all the other categories are lower: –30.4% for mechanical resistance, –28.9% for auxiliary services, –30.7% for losses in the locomotive and the network, and most of all, –57.29% for the energy dissipated in the brakes, with this being the addend with the biggest reduction in consumption.

Different Trains on the Same Line

Andersson and Lukaszewicz compared the energy consumed by a new high-speed train and a conventional train on the same line in Sweden (*8*). They showed that on the same route (from Väterás to Stockholm, Sweden), a conventional train took 78 min and consumed 0.042 kWh per seat kilometer in 1994, whereas in 2004, the X2000 tilting high-speed train, which had one more stop, took 53 min (–32%) and consumed 0.030 kWh per seat kilometer (–28%).

Same Train on Different Lines

The levels of energy consumption of the same train on two lines (a high-speed line and a conventional line) between the same points were previously analyzed and compared (*2*) for two specific cases: an Alvia train between Lleida and Roda de Bará, Spain, and a Talgo train between Córdoba and Antequera, Spain. The Talgo trains (variable track gage) stopped running on the conventional lines and began to operate on the new high-speed lines in December 2006.

In each case, by changing from the conventional line to the high-speed line and running between the same stations, the Alvia and Talgo trains achieved energy consumption reductions in the pantograph of 16% and 8%, respectively. These are compatible with average speed increases of 64% and 78%, respectively.

TECHNICAL REASONS FOR LOWER LEVELS OF ENERGY CONSUMPTION

In all the real and simulated cases analyzed (different trains and lines, different trains on the same line, and the same train on different lines), the results of the analysis coincide with the overall result observed for all Spanish trains: trains that run at higher speeds consume less energy.

As this result clearly contradicts the existing common misconception, it is worth analyzing the reasons for this lower level of energy consumption.

Components of Consumption Function

The dynamics of trains are explained in detail by the author (*9*), who also analyzed the physical and technical reasons for the differences in the levels of energy consumption between high-speed and conventional trains (*3*). The reasons for the differences in the levels of energy consumption are explained by using data from both of those previous studies.

The consumption function is specified by considering the amount of energy released by the train, to which the losses that occur from the moment that the energy leaves the power station to the moment that it leaves the train are added.

Energy Needed to Overcome Mechanical Resistance on Straight Track and Curves

The energy needed to overcome mechanical rolling resistance depends on the specific value of such resistance, the mass, and the distance traveled.

It can be assumed that this resistance is proportional to the train's mass, and the constant of proportionality (a) ranges from 1.2 to 2 decanewtons per metric ton (daN/t) for conventional trains and from 0.5 to 0.9 daN/t for high-speed trains.

On curves, the energy required is proportional to the mass and to the distance traveled and is inversely proportional to the radius of the curve. The author proposed that each line be modeled so that it is defined as an equivalent coefficient of curves (a_c) that would be added to the specific coefficient of the mechanical resistance (*9*). The values of coefficient a_c would vary between 1.424 daN/t (the line from Betanzos to Ferrol, Spain) and 0.08 daN/t (the high-speed line from Madrid to Barcelona).

Energy Needed to Overcome Aerodynamic Drag

The energy needed to overcome aerodynamic drag depends on the square of the instantaneous speed and on the air density. It can be broken down into two addends, one of which is due to pressure (normal forces) and the other of which is due to friction (shear stresses).

Drag due to pressure forces occurs at both the front and the rear of the train and depends [with a proportionality coefficient (c_p)] on the train's cross-sectional surface area (S_t) and on its form [coefficient (C_x)]. It also depends on the air density:

$$E_{rap} = c_p \times S_f \times \int S^2 \times dl$$

where

E_{rap} = energy needed to overcome the aerodynamic pressure drag (kWh),
S_f = cross-section area of the train (m^2), and
dl = differential element of the length.

For conventional trains, the value of c_p is about 0.0022 daN/[(km/h)^2 · m^2], whereas for high-speed trains, the value is closer to 0.00096 daN/[(km/h)^2 · m^2], when a cross-sectional surface area of about 12 m^2 is assumed for both cases (*9, 10*).

Friction drag due to shear stresses occurs on the train's wet surface (S_m). The energy needed to overcome this drag on a route is

$$E_{raf} = c_f \times S_m \times \int S^2 \times dl$$

where E_{raf} is the energy needed to overcome the aerodynamic friction drag (kWh).

For conventional trains, the value of the coefficient c_f (which depends on surface continuity and quality) is about 0.00003 daN/[(km/h)^2 · m^2], whereas for high-speed trains the value is closer to 0.000021 daN/[(km/h)^2 · m^2], if a wet perimeter of 11 m is assumed in both cases (*9*).

Energy Dissipated by the Brake and Not Utilized

On any given route, the train needs additional energy to increase its speed and to climb upgrades. However, instead of being lost, this energy accumulates in the train in the form of kinetic and potential energy, respectively, and can be used to overcome the rolling resistance, in which case it is not wasted.

As the energy needed to overcome rolling resistance has already been considered in the previous addends (whatever the origin may be, i.e., external energy or the energy accumulated in the train itself), it can be understood that the energy lost is only that which is dissipated by the train's brake. This dissipation occurs in two cases: when the train brakes to reduce its speed or when it brakes to descend a steep downgrade without exceeding the maximum speed.

Kinetic Energy Dissipated in Speed Reductions

The kinetic energy dissipated in speed reductions depends on the vehicle's mass, the rotating mass, the distance between technical and commercial stops, and the distance between equivalent stops (*9*).

$$E_{frepar} = \frac{1}{2} \times (M + M_{rot}) \times \left(\frac{1}{D_{pc}} + \frac{1}{D_{pe}} + \frac{1}{D_{pt}} \right) \times (S^2) \times L$$

where

E_{frepar} = energy dissipated by the train's brake because of the speed reductions (kW),
M = mass of the train (metric tons),
M_{rot} = equivalent rotating train mass (metric tons),
D_{pc} = number of stops,
D_{pe} = homogeneity of the speed profile,
D_{pt} = distance between technical (noncommercial) stops, and
L = length (km).

The amount of energy dissipated in the speed reduction depends on M and M_{rot}, D_{pc}, and D_{pe}. D_{pc} and D_{pe} are higher in high-speed trains.

Kinetic Energy Dissipated When Braking on Downgrades with a Higher Value

The balanced downgrade (p_b) is defined as the negative gradient for which a train, without braking or the application of traction, maintains a speed equal to its maximum speed on the line section. The higher the train's authorized speed on the line, the greater the balanced downgrade. If the actual existing downgrade (p_a) is greater than the balanced downgrade ($p_a > p_b$), the train must brake so that it does not exceed the maximum speed.

Therefore, the lower that the balanced downgrade is, the more the train will brake on gradients; that is, the lower the maximum speed, the more the train will brake. Thus, for example, when the authorized speed of the Madrid–Barcelona train changes from 220 to 300 km/h, the energy dissipated by the brake on gradients is reduced by 32%.

Energy Consumed by Auxiliary Services of the Train

A train's auxiliary services are systems that consume energy for technical purposes (compressors, ventilators, etc.) and for the comfort of passengers (heating, air conditioning, lighting, etc.)

The amount of energy consumed by these services is proportional to the time during which they operate; therefore, if the average speed increases, the level of consumption per kilometer decreases in the same proportion as the average speed increases.

The energy consumption of the auxiliary services is not directly related to the speed; thus, for the verification of a typical high-speed case, a 50% increase in the average speed means a 29% reduction in the energy consumed by the auxiliary services.

The auxiliary services continue to consume energy during stops, even though the train does not cover any distance at all, and therefore, an increase in the amount of time at a stop means a higher level of consumption per kilometer. The amount of energy consumed by auxiliary services in high-speed trains is also reduced for this reason, as they make fewer and shorter stops than conventional trains.

Energy Loss Coefficient

Power stations must produce an amount of energy for a train that is equal to the net energy required at the train's wheel rim (or at the inlet of the auxiliary services) plus the loss of any kind that occurs until the electricity is converted into energy that can eventually be used.

The energy lost during the voltage change processes and during the transmission of the energy from the power station to the train is proportional to the energy that reaches the locomotive, and the loss coefficient depends on the train's operating voltage (a higher voltage means fewer losses) and also on the electrification characteristics (the cross section of the conductors, the distance between substations, etc.)

In Spain, high-speed lines are electrified at 25 kV of alternating current (AC) and conventional lines are electrified at 3 kV of direct current (DC). The higher voltage means fewer ohmic losses during the transformation and transmission processes. The additional amounts that must be produced, on top of the amounts consumed by the pantograph, are 22.6% for trains operating at 3 kV of DC and 8.8% for trains operating at 25 kV of AC (*11*). This difference, like the one derived from the use of the regenerative brake, is not inherent to the high-speed system but does occur often.

Shorter Distances Between the Same Points

In all the components of the consumption function, the distance traveled is multiplied by the corresponding monomial. Therefore, it can be said that the energy consumed by a train is proportional to the distance traveled.

High-speed railway layouts imply shorter distances than those of conventional lines between the same points, and the amount of energy consumed is lower by the same proportion, all other factors being equal. Data presented elsewhere show that, in Spain, the average distance of high-speed lines is 12% shorter than that of conventional lines between the same points (*12*).

COMPARISON OF ENERGY CONSUMPTION AND EMISSIONS BETWEEN HIGH-SPEED TRAINS AND OTHER MODES OF TRANSPORTATION

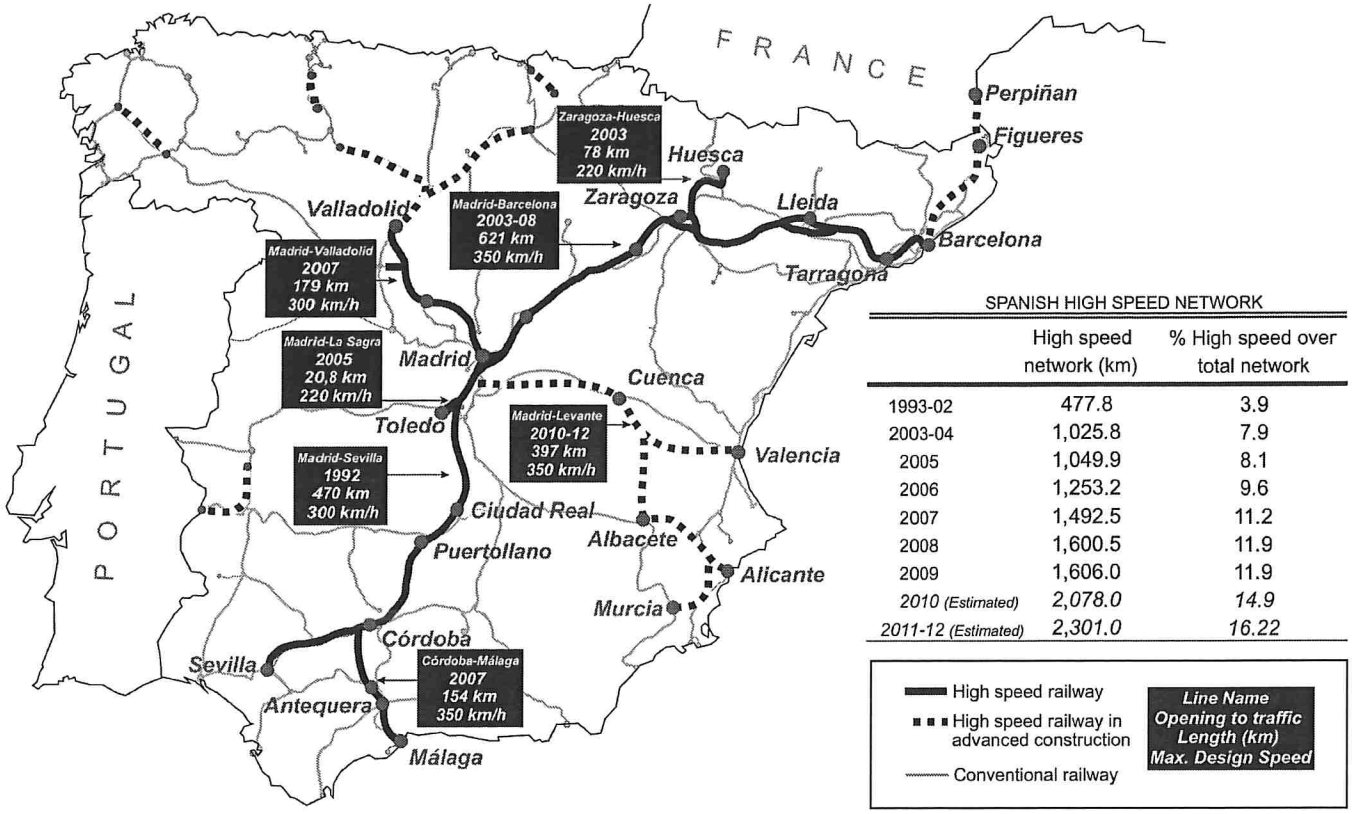
The previous section highlighted the fact that the energy consumption of a high-speed passenger train, under normal operating conditions, is usually lower than that of a conventional passenger train. This section determines the energy efficiency of a high-speed train in comparison with the efficiencies of other modes of passenger transportation under real and comparable conditions. The final objective is to establish the effect of high speed on the total emissions on a corridor.

This analysis for some Spanish high-speed lines has been published elsewhere (*4*). The findings presented in that article indicate that comparison of the amounts of energy consumed by various alternative modes of transportation usually gives rise to numerous and divergent results, depending on the hypotheses used in relation to vehicle occupancy rates, emissions factors, the efficiencies and loss of power generation, the transmission and conversion system used, and the characteristics of the service (speeds, numbers of stops, etc.) (Figure 1).

Numerous studies use average values, sometimes taken from incompatible cases in which the values occasionally correspond to those from different countries (bear in mind that the emissions factors of power generation systems may vary from 1 to 4 from one country to another and from 1 to 1.5 from one year to another in the same country).

Cases Analyzed

In an attempt to overcome the difficulties outlined above, the analysis methodology set out previously was used (*4, 11, 13*), and



SPANISH HIGH SPEED NETWORK

	High speed network (km)	% High speed over total network
1993-02	477.8	3.9
2003-04	1,025.8	7.9
2005	1,049.9	8.1
2006	1,253.2	9.6
2007	1,492.5	11.2
2008	1,600.5	11.9
2009	1,606.0	11.9
2010 (Estimated)	*2,078.0*	*14.9*
2011-12 (Estimated)	*2,301.0*	*16.22*

FIGURE 1 High-speed rail network in Spain.

instead of using average values of unknown origin, the analysis was applied to 10 routes in Spain that can be considered representative of the new high-speed lines and the services provided on them:

- The high-speed line from Madrid to Seville (data were obtained from more than 16 years of operation);
- The high-speed line from Madrid to Barcelona;
- The high-speed line from Madrid to Toledo, which has a typically regional service, no intermediate stops, and a short distance (75 km); and
- Three new high-speed lines [Córdoba to Málaga, Madrid to Valladolid, and Madrid to Levante (the last of which includes five routes)], albeit with the limitations imposed by a lack of sufficient data for high-speed operations.

General Methodological Comments

Some of the key or differential aspects of the methodology used are explained below [additional details are provided elsewhere (*4*)].

Distances

An important element of the analysis is that it considers the actual distances traveled by each mode, which vary significantly on the same route. By taking access into account, the difference in the distance between the mode with the longest distance and the one with the shortest is –33%, in the case of Madrid to Valencia, and –30%, in the case from Madrid to Valladolid. The smallest difference corresponds to the route from Madrid to Alicante (–13%).

Load Factor

One of the key points in modal comparison studies is the load factor considered, because the energy consumed by and the emissions required to move a vehicle depend little on whether the vehicle is full or empty. Therefore, the attribution of energy consumption to passengers gives rise to significant variations, depending on how many passengers are being transported. As average load factors usually vary considerably depending on the mode of transportation, this factor has a significant influence on the results of the intermodal comparison. The analysis presented here used as averages the actual load factors (known or estimated) for each mode of transportation and route in Spain.

Results of Comparison Between Modes

The comparison of the CO_2 emissions of the various modes of transportation shows that on 7 of the 10 routes analyzed, the high-speed train produces fewer emissions than any other mode. On the other three routes, the conventional train has the lowest value. On average, the emissions of the conventional train are 41.9% higher than those of the high-speed train (37.9 and 26.7 g of CO_2 per passenger straight kilometer).

On the routes from Madrid to Valencia, Albacete, and Alicante, the conventional train has lower levels of energy consumption. This is mainly due to the longer distances traveled by the high-speed train.

TABLE 1 Greenhouse Gas Emissions by High-Speed Trains Compared with Those of Other Modes of Transportation in Spain

	Straight Distance (km)	Emissions ($kgCO_2$)/Pass × km Average Load Factor					Emissions (times those of HST)				
		Car	Coach	Plane	Conv. Train	HS Train	Car	Coach	Plane	Conv. Train	HS Train
Madrid–Barcelona	486	129.7	29.9	145.9	35.3	28.4	4.56	1.05	5.13	1.24	1
Madrid–Málaga	417	132.2	32.4	146.7	28.8	22.1	5.97	1.46	6.63	1.3	1
Madrid–Sevilla	396	136.6	35.0	154.9	32.1	23.7	5.76	1.48	6.54	1.35	1
Madrid–Alicante	359	122.5	33.3	163.3	28.8	32.9	3.73	1.01	4.97	0.88	1
Madrid–Murcia	349	116.5	29.9	195.1	40.3	27.2	4.28	1.10	7.18	1.48	1
Madrid–Valencia	302	120.2	33.2	208.7	29.2	33.7	3.56	0.99	6.19	0.87	1
Madrid–Albacete	224	121.6	32.6	0.0	21.7	22.6	5.38	1.44	0.00	0.96	1
Madrid–Valladolid	162	87.6	35.9	0.0	39.8	25.4	3.45	1.42	0.00	1.57	1
Madrid–Cuenca	139	80.0	33.1	0.0	95.0	27.9	2.87	1.19	0.00	1.57	1
Madrid–Toledo	70	83.2	28.0	0.0	28.4	23.1	3.61	1.22	0.00	3.41	1
Average (GCO_2/passenger .km)		113.0	32.3	169.1	37.9	26.7	4.23	1.21	0.00	1.42	1
		Primary Energy (kWh)/Pass × km Average Load Factor					Primary Energy (times those of HST)				
Average (kWh/passenger .km)		0.48	0.12	0.54	0.26	0.19	2.48	0.65	2.79	1.37	1

NOTE: Load factors are 0.3 to 0.4 for cars, 0.55 to 0.61 for coaches, 0.75 for planes, 0.64 for long-distance conventional trains, 0.35 for regional trains, 0.70 for long-distance high-speed trains, and 0.56 for regional high-speed trains. Conv. = conventional; HST = high-speed train; HS = high-speed.
SOURCE: Author's study (*4*) was expanded for this analysis.

The conventional train and the bus share second place in lower emissions on the routes on which AVE has the lowest emissions: on the Málaga and Seville routes, the conventional train emits less than the bus; and on the Barcelona, Valladolid, Murcia, and Cuenca routes, the bus emits less (it should be pointed out that the last two routes are the only ones on which the conventional train runs on diesel). On the Madrid–Toledo route, they emit practically the same, although the differences are minimal in all cases. Overall, the bus emissions are 19.6% higher than those of the high-speed train.

Overall, car emissions are 4.5 times higher than those of the high-speed train and 5.1 times higher than those of the plane (on the routes on which air services exist) (Table 1).

EFFECT ON A ROUTE

The consumption and emission differences between the various modes of transportation on a route are, as can be seen, quite considerable. Furthermore, whenever a high-speed train establishes itself on any route, it invariably gives rise to significant changes in the modal distribution and the size of the market (*14*).

Specifically, whenever a high-speed train coexists with other modes of transportation, it attracts passengers away from planes, private cars, buses, and conventional trains (regardless of whether conventional trains continue to exist or disappear).

The high-speed train offers an economical cost–time combination that makes the cost lower than the cost of all preexisting modes for certain market segments, thereby inducing journeys that previously did not exist.

In this way, determination of the energy (and emissions) balance when a high-speed train is present on a route cannot be limited to a comparison of the energy consumption of the high-speed train with that of the conventional train. In fact, in practically every case, passengers coming from the conventional train represent only a small proportion of high-speed train passengers.

The balance should therefore take into account all passengers who have diverted from one mode of transportation to another, and the energy consumption and emissions differentials should be applied to the number of passengers of each modal transfer. It is also necessary to add the levels of energy consumption and emissions of the high-speed train that are attributable to the increase in overall demand on the route, that is, to the induced demand, which, despite being a positive factor in terms of mobility, requires the consumption of larger amounts of energy.

To be able to carry out this analysis, the levels of consumption and emissions on each origin–destination route for each mode of transportation, as well as the intermodal transfer rates that result from the appearance of passenger high-speed rail, need to be known.

The energy consumption and emission figures for numerous routes in Spain can be obtained from Table 1. For other lines in the Spanish sphere, this information can be obtained by interpolating the data to similar cases, as will be explained below.

It is difficult to ascertain intermodal transfer rates and the induced demand a priori, because there are many different types of cases that depend on the length of the route; on whether air services exist and, if they do, their characteristics; on the road transportation options; and, in short, on the characteristics of the high-speed rail services. However, it also seems possible to use certain patterns to make a general approximation for each type of case, although this should not prevent an attempt to investigate the data for each specific route from being made.

Madrid–Seville Route

In 2005, a total of 2.5 million passengers were transported on high-speed trains between Madrid and Seville with an average journey time of 2 h 28 min (at an average speed of 190.7 km/h).

The customer surveys conducted during the early years of operation of the AVE included the following question: "If the AVE had not

existed, which mode of transportation would you have chosen for this journey?" The possible responses were "conventional train," "plane," "car," "coach," or "I wouldn't have made the journey." In response, 13% of the AVE users said they would have traveled by conventional train, 26% by plane, 24% by private car, and 3% by coach, whereas 34% said they would not have made the journey.

The data used in the present study are based on an evaluation of the variation in CO_2 emissions on the Madrid–Seville route if the AVE did not exist. For this purpose, the high-speed train passengers in 2005 were spread among the other modes of transportation (or it is assumed that they would not travel), according to the percentages deduced from the survey. In the transfer to each mode, there is a difference in emissions (always with higher emissions in the event that there is no AVE, except in the case of the induced demand, which, logically, would produce fewer emissions if the AVE did not exist).

Transportation of the 2.5 million passengers who use the AVE represents an annual CO_2 emission total of 23.7 million kilograms (Mkg) of CO_2, but if the AVE had not existed, the emissions would have amounted to 78.04 Mkg of CO_2, which means that the emissions would have multiplied by 3.3. On average, each AVE passenger prevents the emission of 21.6 kg of CO_2.

It is necessary to draw attention to the fact that the percentage of travelers who would travel by conventional train in the absence of the AVE is so low (only 13% of AVE passengers would choose the conventional train as an alternative) that the emissions reductions due to the AVE's lower level of consumption in comparison with the level of consumption of the conventional train are insignificant (1.09 Mkg of CO_2 compared with a total reduction of 54.37 Mkg of CO_2). This shows that the main advantage of the AVE is not that it consumes less energy and emits fewer greenhouse gases than the conventional train. The main advantage is that, thanks to its speed, it is capable of attracting a high percentage of travelers away from energy-inefficient modes of transportation, such as aircraft and private cars.

Madrid–Barcelona Route

The case of the Madrid–Barcelona route, on which the AVE has been operating since February 2008 with a minimum journey time of 2 h 38 min, is substantially different from the previous case (Madrid–Seville route) for various reasons: a longer distance (+31.8%) and, therefore, a longer AVE journey time; a more mature market with greater mobility previously induced by the most traveled air shuttle service in the world; the availability of frequent air transportation services; and the presence of low-cost airlines whose prices are lower than those of the AVE. All of these characteristics suggest that the results for the Madrid–Barcelona route would be significantly different from those for the Madrid–Seville route.

Because of its importance, countless studies have dealt with the previous and anticipated traffic levels on the Madrid–Barcelona route. Of all those studies, data from the study of López Pita and Robusté (*15*) were selected. Use of those data, treated in a way similar to the way in which the data for the Madrid–Seville route were treated and by use of the unit emission data presented above, would lead to the CO_2 balance shown in Table 2 (*4, 15*).

From these results it can be deduced that the carbon dioxide emitted by the 3.48 million travelers who use the AVE in 1 year would be multiplied by 3.9 if there were no AVE. The existence of AVE produces an annual emissions reduction of 140 Mkg of CO_2, with an average reduction per AVE passenger of 40.3 kg of CO_2.

By comparing both cases, it can be deduced that on a route on which the plane is strongly established, as is the case for the Madrid–Barcelona route, the high-speed train attracts more travelers away from the plane (in absolute numbers), although the induced demand is lower (10%). The net effect is that the reduction in emissions (and, therefore, in energy consumption) on the Madrid–Barcelona route is even higher than that on the Madrid–Seville route.

Although the high-speed train emits (and consumes) less than the conventional train, the overall savings of all the modes of transportation on the route due to the existence of the AVE is between 2 and 15 times greater than the savings produced by comparing the AVE with the conventional train. This again underlines the earlier affirmation about the advantage being not only the lower levels of energy consumption by the AVE in relation to those of the conventional train but also, and above all, the leverage effect on the route owing to transfers from other modes of transportation.

There is a relationship between the emissions avoided on the route per AVE passenger kilometer and the average speed on the route in

TABLE 2 Difference in CO_2 Emissions on Madrid–Barcelona Route With and Without High-Speed Train (HST)

HST: 2 h 38 min 621 km	kg CO_2/Passenger Madrid–Barcelona	M Passengers with HST	Emissions with HST (Mkg CO_2)	M Passengers Without HST	Emissions Without HST (Mkg CO_2)	Difference Without HST (kg CO_2/v)	Difference Without HST (Mkg CO_2)
High-speed train	13.81	3.48	48.1	0.00	0.00	0.00	0.00
Conventional train	17.14	0		0.57	9.77	3.33	1.90
Plane	70.89	0		2.20	155.95	57.07	125.56
Car	63.05	0		0.36	22.70	49.24	17.73
Bus	14.51	0		0.00	0.00	0.70	0.00
New passengers	0	0		0.35	0.00	–13.81	–4.84
Total		3.48	48.1	3.48	188.42		140.35

NOTE: HST's average speed = 236 km/h. If there were no HST, the emission would be multiplied by 3.919. The 58% of HST passengers who use both train and plane avoid 40.33 kg of CO_2 per passenger. M = mega (10^6), v = passenger.
SOURCE: Emissions figures are from the author (*4*), and the numbers of passengers by means are from López Pita and Robusté (*15*).

question. The amount of emissions avoided by the high-speed train increases as the train's speed increases both on routes with air services and on routes with no air services. On the routes with no air services, the emissions avoided by the AVE are lower, although they grow more rapidly as the train's speed increases, at least in the cases analyzed.

CONCLUSIONS

The following conclusions can be drawn from the findings presented in this paper:

1. The high-speed train, under its normal operating conditions, consumes less energy and produces fewer emissions (29% less, on average) per passenger transported than a conventional train running between the same points at a lower speed.
2. The reasons for the high-speed train's lower levels of energy consumption are to be found in certain intrinsic features of the high-speed rail system (more homogeneous speed profile, fewer stops, fewer curves, higher power supply voltages, etc.). Certain consumption addends (such as auxiliary services) are also directly reduced with speed. Moreover, high-speed lines usually cover shorter distances than conventional lines for the same routes.
3. In terms of energy consumption and emissions, the great advantage of the high-speed train is that it uses electric traction, which means fewer greenhouse gas emissions and a lower contribution to the exhaustion of fossil fuels than other modes of transportation. It shares this advantage with the conventional electric train.
4. Precisely because of its speed, the high-speed train is capable of attracting a significant proportion of travelers away from the airplane and the private car. Therefore, the main advantage of introducing a high-speed rail line does not stem from replacing the conventional train: typically, the high-speed train prevents the emission of 3 kg of CO_2 per passenger in relation to the conventional train, whereas on the route as a whole, the introduction of high speed has a leverage effect that prevents the emission of 31 kg of CO_2 per high-speed rail passenger.
5. High-speed rail lines must be designed to obtain a high level of energy efficiency (a uniform speed profile, adequate gradients, short distances, etc.), but above all, they must be capable of offering reduced journey times to capture plane passengers.

ACKNOWLEDGMENTS

Helpful comments and suggestions were provided by Victor Aragonés, Scott Greene, and Melissa Shurland.

REFERENCES

1. Minayo, F., and A. García. *Relación Entre la Alta Velocidad Ferroviaria y los Costes Operativos no Energéticos.* Fundación Caminos de Hierro, Madrid, Spain, 2008.
2. García Álvarez, A. El Tren de Alta Velocidad No Es un Depredador de Energía. *Dyna,* Vol. 80, No. 5, June 2005, pp. 33–38.
3. García Álvarez, A. Incidencia del Tren de Alta Velocidad en el Consumo Energético y Emisiones del Sector Transporte. Presented at 6th Science Week, Madrid, Spain, Nov. 17, 2006.
4. García Álvarez, A. Consumo de Energía y Emisiones del Tren de Alta Velocidad en Comparación con Otros Modos. *Anales de Mecánica y Electricidad,* Vol. 84, Fascicle V, Sept.–Oct. 2007.
5. Kemp, R. Take the Car and Save the Planet. *Power Engineer,* Vol. 18, No. 5, 2005, pp. 12–17.
6. Kemp, R. J. The European High Speed Network. In *Passenger Transport After 2000 AD,* Feilden, Wickens, and Yates, London, 1993.
7. *Memoria Renfe 2005.* Dirección de Comunicación y Relaciones Externas, Madrid, Spain, 2006.
8. Andersson, E., and P. Lukaszewicz. *Energy Consumption and Related Air Pollution for Scandinavian Electric Passenger Trains.* Report KTH/AVE 46. Department of Aeronautical and Vehicle Engineering, Royal Institute of Technology, Stockholm, Sweden, 2006.
9. García Álvarez, A. *Dinámica de los Trenes en Alta Velocidad.* Fundación de los Ferrocarriles Españoles, Madrid, Spain, 2004.
10. Bernhard, M., and M. Guieu. Mesures Récentes de la Résistance a l'Avancement de Matériels Roulants. *Revue Générale des Chemins de Fer,* April 1976, Paris.
11. Pilo, E. Diseño del Sistema de Alimentación Eléctrica. *Jornadas de Eficiencia Energética en el Ferrocarril,* March 2006, p. 21.
12. García Álvarez, A., and E. Fernández González. Recorridos y Cociente Entre Trayectoria y Desplazamiento en el Transporte por Ferrocarril. In *Notas Técnicas Enertrans 9.* Fundación de los Ferrocarriles Españoles, Madrid, Spain, Jan. 2008.
13. García Álvarez, A. Normalización de los Consumos Energéticos de los Trenes de Viajeros. Presented at 3rd Railway Innovation Conference, Tenerife, Spain, 2007.
14. Barrón de Angoiti, I. Introducción a la Alta Velocidad Ferroviaria. In *Apuntes del Master de Sistemas Ferroviarios,* 6th ed. Instituto Católico de Artes e Industrias, Madrid, Spain, 2007.
15. López Pita, A., and F. Robusté. Impact of High-Speed Lines in Relation to Very High Frequency Air Services. *Journal of Public Transportation,* Vol. 8. No. 2, 2005, pp. 17–35.

The Passenger Rail Equipment and Systems Integration Committee peer-reviewed this paper.

Genetic Algorithm–Based Column Generation Approach to Passenger Rail Crew Scheduling

Mindy Liu, Ali Haghani, and Shahabeddin Toobaie

The goal of the present study was to develop and apply a genetic algorithm–based column generation heuristic to solve a passenger rail crew scheduling problem in North America. The crew scheduling problem minimized the total cost of payment to crew members on the basis of the number of hours on board, the number of hours held away from a crew base, the number of nights of lodging, and the number of onboard and away meals. Payment regulations also dictated an overtime payment and a guaranteed salary per week. Additional problem constraints included restrictions on the maximum number of continuous working hours, the maximum number of days worked per week, and the minimum number of hours of rest. The proposed heuristic produced solutions with improvements of total cost ranging from 3.0% to 27.9%.

The crew scheduling problem (CSP) is the problem of assigning work to crew members to create a minimal-cost work schedule. The purpose of the study described here was to apply a genetic algorithm–based column generation heuristic to solve the passenger rail CSP for Amtrak's Northeast Corridor. For this study, the CSP focused solely on the train and engine (T&E) crew, which consists of conductors, assistant conductors, engineers, and firemen (assistant engineers). The CSP for the Northeast Corridor contained 4,338 total work duties that must be covered by these crew members.

The objective of CSP was to minimize the cost of payment to the crew on the basis of the number of hours worked on board a train, the number of hours held away, the number of nights of lodging provided, and the number of onboard and away meals provided. The constraints of this problem were the work regulations and the payment regulations defined by Amtrak. These constraints set minimum rest times, the maximum number of working days per week, overtime rates, and guaranteed salary. The complexity of crew scheduling constraints also contributed to the complexity of the problem.

TERMINOLOGY

Because the crew scheduling terminology is not standard throughout the literature, the definitions of the relevant terminology follow:

M. Liu and A. Haghani, Department of Civil and Environmental Engineering, University of Maryland, 1173 Glenn L. Martin Hall, College Park, MD 20742. S. Toobaie, Marketing and Product Management Department, National Railroad Passenger Corporation (Amtrak), 60 Massachusetts Avenue NE, Washington, DC 20002. Corresponding author: A. Haghani, haghani@umd.edu.

Transportation Research Record: Journal of the Transportation Research Board, No. 2159, Transportation Research Board of the National Academies, Washington, D.C., 2010, pp. 36–43.
DOI: 10.3141/2159-05

- Duty: a duty corresponds to a single train trip or 1 day of yard work or extra board.
- Extra board: extra board duties are when crew members are on backup duty for other crew. These crew members are not on duty unless unforeseen circumstances force them to be called to work. Crew members on extra board are paid regardless of whether they are called to work to meet the guaranteed minimum weekly salary.
- Shift: one or many duties make up a shift. A shift is a chain of duties that begins and ends at a crew base.
- Crew base: the crew base is the location at which a crew member goes on and off duty and is usually near the crew member's residence.
- Pairing: one or many shifts or a sequence of trips that starts and ends at a single crew base for a given time horizon makes up a pairing. In this study, a pairing consists of all duties for a single crew member over a 1-week period.
- Away: during a shift, a crew member may travel to an away location, a location that is not the crew base and in which a crew change is permitted. When the crew is at an away location, the crew may have rest time and may be compensated for meals and lodging, depending on the length of the away time.
- Deadhead: when a crew deadheads a duty, the crew is not on duty during that train trip. Deadheading may be required to transport a crew to a location to start a shift or to transport a crew back to the crew base at the end of a shift.

PROBLEM DESCRIPTION

The study described here examined the passenger rail CSP for Amtrak's Northeast Corridor, the busiest corridor in the Amtrak network (Figure 1). This corridor covers 32 stations and 16 crew bases and allows for billions of unique schedules. The six problems (referred to as EFN, ACN, EFS, ACS, EF, and AC) are shown in Table 1.

The problem consists of four crew types: conductors, assistant conductors, engineers, and assistant engineers. Conductors and assistant conductors share responsibilities, as do engineers and assistant engineers, which conveniently allows the route information to be divided between those two groups for data processing and solving. With a partition of the single, full network problem as two independent problems, the problem size is dramatically reduced.

The schedule data set lists all current work schedules for each individual crew member. For each individual crew member's schedule, the data indicate the duties, pairings, and away time for a single week. To further reduce the complexity in coding required to process this large amount of data efficiently, the Northeast Corridor zone was partitioned into a north region and a south region because the current train

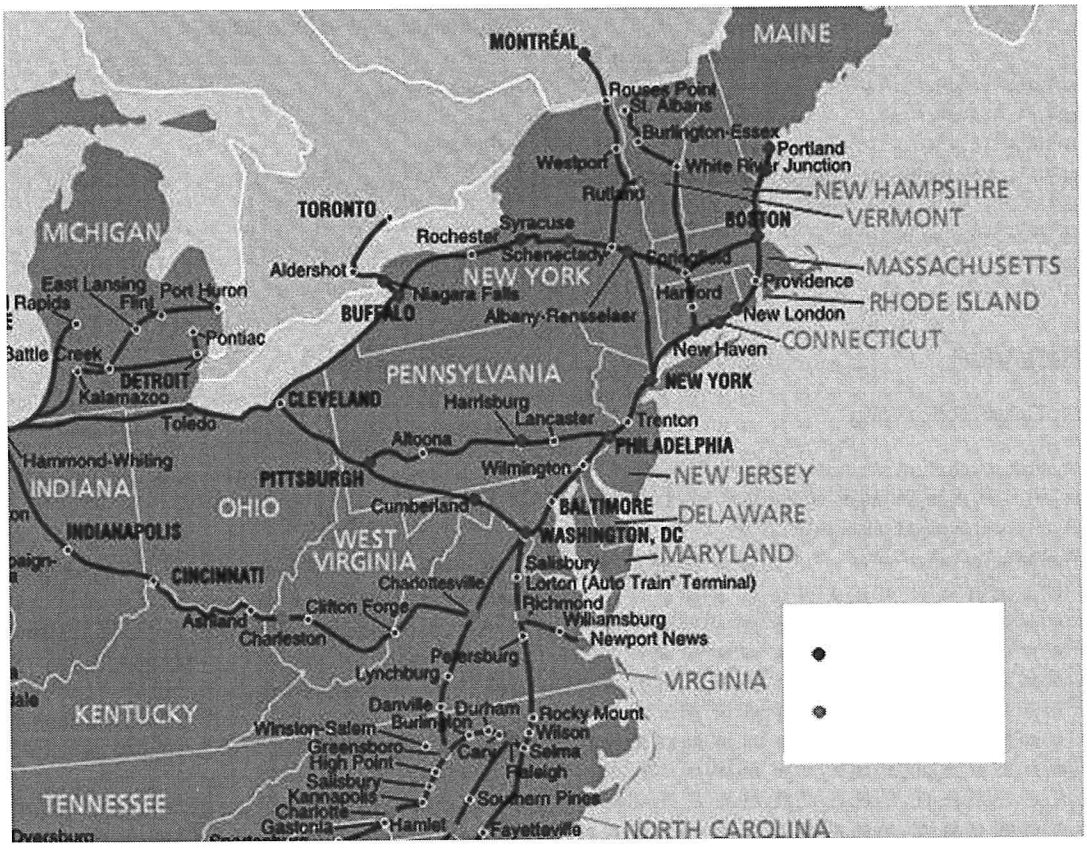


FIGURE 1 Amtrak's Northeast Corridor.

network has a single station at the Washington, D.C., crew base that divides the network into two sections (Table 1). This partitioning, in conjunction with the previously described specialized crew types, creates separate problems for the Northeast.

PROBLEM CONSTRAINTS

Amtrak T&E crew members are subject to specific work regulations. First, T&E crew may be required to stay at away locations between trips or overnight but must return home at the end of a shift.

TABLE 1 Summary of CSPs

Problem	Crew Type	Region	Code	Number of Duties
1	Engineers & firemen	North	EFN	1,044
2	Conductors & assistant conductors	North	ACN	1,178
3	Engineers & firemen	South	EFS	1,049
4	Conductors & assistant conductors	South	ACS	1,067
5	Engineers & firemen	North & south	EF	2,093
6	Conductors & assistant conductors	North & south	AC	2,245

Second, T&E crew schedules permit deadheading of crews. Deadheading may require additional cost for transporting the crew to the appropriate locations but is necessary for some cases to meet all work requirements.

In accordance with federal regulations, the T&E crew is guaranteed minimum rest times. Before a crew works a shift, the crew must be allowed at least a consecutive 8 h of rest during the 24 h before they work. After being on duty for a consecutive 12 h or more, the crew must have at least a consecutive 10 h of rest. In addition, crew members must have at least 1 entire day without work per week.

The regular hourly rate paid to the crew is assumed to be $33.60, which is based on the real Amtrak crew cost. The 1.5 overtime rate is thus equal to $50.40. The categories of payment regulations include shift work payment, meal allowance, lodging allowance, and a guarantee of 40 h of work per week. There are four payment types:

1. Shift work payment. Shift work payment is pay for hours on duty and billable away hours. The on-duty hours are time spent on a train trip, yard work, or extra board hours. When they are held away, crew are not paid for the first consecutive 12 h they are away. After they are held away for a consecutive 12 h, Amtrak pays the crew for up to 8 h every consecutive 24 h they are away. The calculation of shift payment is based on entire shifts rather than individual work duties. The Amtrak regulation states that the first 8 h of a shift are paid at the regular hourly rate and any additional hours in the shift above 8 h are paid at the overtime rate.
2. Meal allowance. If a train duty is at least 5 h long, Amtrak provides crew members an allowance of $3.50 every 5 h. When crew

members are held away for at least 4 h, Amtrak provides an allowance of $6.00 and another $6.00 once every 8 h afterwards.

3. Lodging allowance. Lodging allowance is assumed to be $70.00, and additional lodging allowance is provided once every 24 h after the first 4 h of away time.

4. A 40 h guarantee. If the total working time and paid away time is less than 40 h, Amtrak pays for the additional hours at the regular rate of $33.60/h. If the total working time and paid away time is over 40 h, however, Amtrak pays for hours over 40 h at the 1.5 overtime rate.

RELEVANT RESEARCH

Early Work on Crew Scheduling

Early research on crew scheduling was primarily applied in the transit and air industries (*1–3*). A comprehensive summary of early CSPs and solution methods is provided by Bodin et al. (*4*).

Most of the research on rail crew scheduling in North America has focused on the freight rail industry rather than passenger rail. Although the lessons learned from research on CSPs in either industry may be applicable to other industries, there are also important differences between freight and passenger rail crew scheduling. First, the North American passenger rail industry has special payment rules with guaranteed weekly or biweekly salaries. Amtrak has a unique guaranteed payment rule not found in many other industries. Second, the passenger rail crew payment is based on shifts rather than individual duties. These shorter trips may result in more time away from a crew member's crew base, because more trips are required to fill a single crew member's weekly or biweekly schedule. A third difference between freight and passenger rail crew scheduling is that passenger rail crew normally have set work duties. These differences between freight and passenger rail crew scheduling highlight the need for more research on North American passenger rail CSPs.

At the time of this work, there has been limited research on the passenger rail CSP in the United States. The large body of research done in Europe may have been a response to the deregulation and privatization of the rail industry in European countries. In addition, there has been research in other areas that have larger markets for rail, such as Asia and Australia. This work contributes to the limited body of research on North American passenger rail CSPs.

Set Partitioning Problem Formulation

The method used in this study was the popular set partitioning problem approach. This method was selected over other approaches, such as a network flow formulation or meta-heuristics, because it is able to capture the problem constraints and produce a solution within the constraints of reasonable resources. In the set partitioning problem formulation of the CSP, the decision variable, x_j, is a binary integer variable that represents whether a pairing is selected as a work duty for one crew member. The constraints consist of a matrix of binary values, a_{ij}, that indicates if a pairing j covers a work duty i. Each row in this matrix shows which pairings cover a single work duty. Each column corresponds to one possible pairing or the work for an individual crew member over the defined time horizon. The set partitioning problem models the CSP as a problem of finding a minimum-cost subset of pairings that exactly covers the work requirements. The set partitioning problem formulation is

$$\text{minimize } \sum\nolimits_j c_j x_j$$

subject to

$$\sum\nolimits_i \alpha_{ij} x_j = 1$$

$$x_{ij} \in \{0,1\}$$

where

c_j = cost of pairing,

$$x_j = \begin{cases} 1 & \text{if pairing } j \text{ is selected} \\ 0 & \text{otherwise, and} \end{cases}$$

$$\alpha_{ij} = \begin{cases} 1 & \text{if pairing } j \text{ covers work requirement } i \\ 0 & \text{otherwise.} \end{cases}$$

A major difficulty with the set partitioning problem formulation of the CSP is that of determining all possible pairings for these models. In the full-set partitioning problem formulation, the matrix of constraints contains columns for every possible pairing. For larger CSPs, this results in an extremely large number of columns. For a CSP with thousands of trips, there can easily be an unmanageably large number of possible pairings. In this regard, the first problem is the time-consuming task of enumerating all these possible pairings. In the rail CSP, complex work and payment rules make it difficult and inefficient to determine the set of all possible and legal pairings. The second problem is that even if all pairings can be enumerated, the resulting set partitioning problem, which is a binary integer program, is NP-hard (nondeterministic polynomial-time hard).

Researchers have used various relaxation methods, heuristics, column generation, and branching strategies to solve these problems. Marsten and Shepardson reviewed earlier work in which the set partitioning problem model of the CSP was used (*5*).

To reduce the initial problem size, researchers have traditionally used heuristics (*6–9*) and bounding techniques (*10*). However, for very large CSPs, many researchers have adopted column generation techniques (*11, 12*).

Column generation has also been combined with pricing strategies in the branch-and-price method (*13–15*). Similar to the branch-and-bound method for integer programming problems, the branch-and-price method uses a branching tree to solve for integer solutions. Constraint branching rules are used to solve a pricing subproblem at each node in the branch-and-bound tree. The branch-and-price method involves the generation of new pairings, or columns while an integer program (IP) is solved, whereas the column generation heuristic method generates new pairings in between solving linear program (LP) relaxation problems. Branch and cut is a similar method that involves branch and bound with cutting planes to solve large-integer programming problems (*16–18*).

SOLUTION APPROACH

The problem is formulated as the general set partitioning problem described above. The initial matrix represents the existing Amtrak schedule obtained in fall 2007. To reduce the problem size, a genetic

algorithm is applied to drive the column generation process and to find solutions with reasonable time and computer memory resources.

Genetic algorithms use evolutionary algorithms based on genetic evolution in biology to search for solutions to optimization problems. Although other heuristics have been used in conjunction with column generation for passenger rail CSPs in the past, genetic algorithms have not been used for this application. The solution approach is outlined in Figure 2.

In large CSPs, it is not efficient to enumerate all possible pairings or columns of the master problem, so a restricted subset of the set of all feasible pairings is created. New pairings are created by the genetic algorithm, which uses LP duality theory to decide which new pairings are "good." According to LP duality theory, a solution improves when variables with negative reduced costs enter the basis. The new pairings are priced out to determine which pairings improve the LP relaxation of the master problem and, thus, the master problem. By using the dual variables from the LP solution, the reduced cost of each pairing can be calculated by using the following equation from LP duality theory:

$$\bar{c}_j = c_j - \sum_i \lambda_i \alpha_{ij}$$

where

$\bar{c}_j$ = reduced cost of pairing j,
c_j = cost of pairing j,
λ_i = dual variable of duty i, and

$$a_{ij} = \begin{cases} 1 & \text{if pairing } j \text{ covers work requirement } i \\ 0 & \text{otherwise.} \end{cases}$$

In each iteration, the restricted LP relaxation of the master problem is solved by calling the commercial optimization package software CPLEX through the CPLEX callable library. The purpose of solving the restricted LPs is to find the dual variables, or shadow prices, corresponding to the duties in the problem for each restricted subset of pairings. These values are used to calculate the reduced cost of new pairings in the pricing subproblem. Thus, the objective of this pricing subproblem is to minimize the reduced cost of pairings not already in the restricted subset. If the solution to this pricing subproblem is negative or zero, then the corresponding pairing should enter the restricted subset. To make the iterations more efficient, many pairings with negative reduced costs can be brought into the restricted subset in each iteration.

In the genetic algorithm, a pairing is stored as a binary chromosome corresponding to a column in the binary matrix of the set partitioning problem formulation of the CSP with the addition of the crew base. The fitness of the pairing is equal to the reduced cost of the pairing, which is calculated from the cost, dual variable, and assigned duties. The genetic algorithm initializes with the current subset of pairings in the restricted problem. The two pairings with the lowest fitness from the population are selected to be "parents." Once parents are selected, the parents reproduce to create "children." Reproduction occurs through a random, one-point crossover. The one-point crossover splits each parent into two parts and retains the first part of a parent and the second part of the other parent. The point at which the crossover occurs is selected by a random number generator in each iteration. Crew bases are stored at the beginning of a chromosome, and as a result, a new pairing retains the crew base of the first parent.

The crossover creates a possible pairing, and a subsequent feasibility operator tests to ensure that the pairing is feasible. A pairing is considered feasible if it is feasible to the overall CSP and it is not already in the subset of pairings in the restricted problem. In the application of the genetic algorithm to the CSP in this study, a large

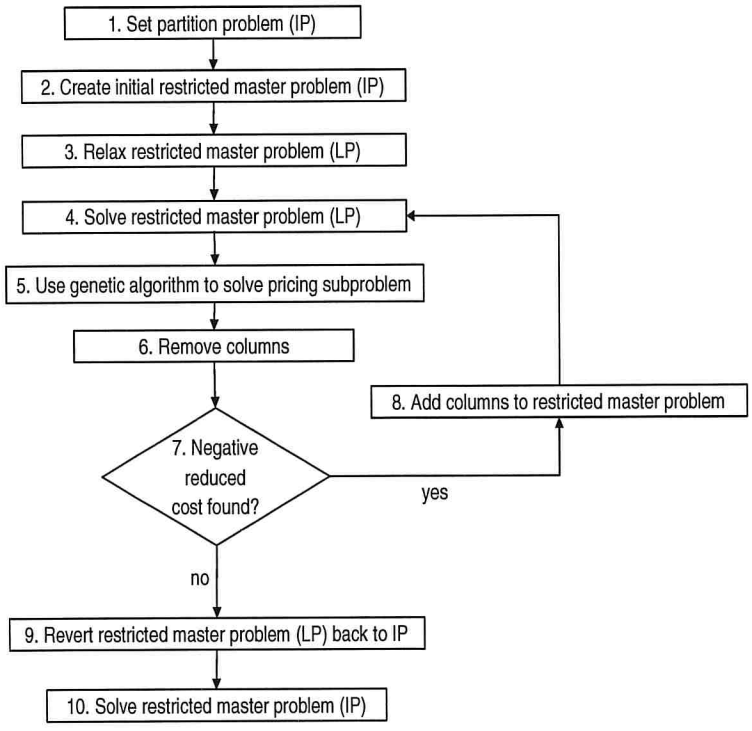

1. Set partition problem (IP)
2. Create initial restricted master problem (IP)
3. Relax restricted master problem (LP)
4. Solve restricted master problem (LP)
5. Use genetic algorithm to solve pricing subproblem
6. Remove columns
7. Negative reduced cost found?
yes
8. Add columns to restricted master problem
no
9. Revert restricted master problem (LP) back to IP
10. Solve restricted master problem (IP)

FIGURE 2 Solution approach.

number of pairings are infeasible because of the constraints. An additional mutation operator facilitates the genetic diversity of subsequent populations. This reproduction process iterates to create many new populations of pairings for the restricted problem.

The genetic algorithm terminates when a fixed number of feasible pairings is created. The population size, *n,* is limited to a defined number to prevent the problem from becoming excessively large. These pairings are added to the restricted problem. The genetic algorithm also terminates if no new feasible pairings are found within a defined amount of time. The latter stopping criterion prevents searching for an excessive amount of time.

The pairing generation process creates a large number of new pairings. To prevent an excessive number pairings in the pricing subproblem, pairings that are not selected to be in the restricted LP solution and that have a positive reduced cost are removed. These pairings are selected to be removed because these are less likely to improve a subsequent LP solution.

When no further pairings with a negative reduced cost can be found by the genetic algorithm, the relaxed, restricted master problem can be solved to optimality or near optimality. The final relaxed, restricted master problem is reverted back to an IP. This IP is then solved to optimality by using CPLEX. The resulting solution from this IP is the final solution to the CSP.

TESTING STRATEGY

To test the proposed heuristic solution approach, three reduced-size problems were generated on the basis of the actual data. The problems were scaled versions of the real-world problems and had 26, 28, and 30 duties each. These problems were solved to optimality by using the traditional set partitioning problem formulation in CPLEX and the proposed heuristic.

The real-world problems were then solved by using the proposed heuristic approach. Sensitivity analysis was then conducted to determine the effects of changes in cost, schedule, and population size on the solution. Each of the four large problems, ACN, EFN, ACS, and EFS, was solved by using the heuristic with population sizes *n* of 2,000, 4,000, and 20,000 pairings, whereas the larger AC and EF problems were solved with a population size of 2,000.

RESULTS

Generated Problems

The heuristic solved the smaller of the two problems generated to optimality in less time than the traditional set partitioning problem method. The heuristic, however, did not solve the largest generated problem with 30 duties to optimality. The heuristic solution improvement was 19.9%, whereas the optimal solution improvement was 20.3%. The total cost of the final heuristic solution was within 0.5% of the optimal solution. The proposed heuristic also reached a final solution in 31 s, whereas 6,284 s was required for the optimal solution by the traditional set partitioning problem formulation method. Heuristic and optimal solutions for the generated problems are shown in Table 2.

Partitioned Problems

The best solution for the ACN network was found for a population size of 20,000. This parameter produced a crew schedule with a 15.8% cost reduction and an annual savings of $3,516,344. The results for the EFN, ACS, and EFS problems were similar to those for the ACN problem. A summary of the total improvements is provided in Table 3.

Discussion of Results

The results from the problems generated indicated that the heuristic calculation time increased less than the traditional set partitioning problem calculation time as the number of duties increased. The excessive calculation time required to find optimal solutions for the real-world-sized problems demonstrates the need for a more efficient solution method. In this work, for the problems with more than 1,000 duties, the generation of all feasible duties and then solution of the problem to optimality by use of the traditional method become nearly impossible because of computer memory and calculation time constraints.

The longer calculation time for the EFN problem compared with that for the ACN problem suggests that because more pairings need to be generated and stored, the heuristic also finds many more infeasible pairings. An examination of the pairings created during the solution process confirms that many more infeasible pairings are created and removed when the population size limit is larger. This long calculation time would not be appropriate for short-term crew scheduling but may still be acceptable for long-range scheduling. The memory usage required for calculation of the solution for this EFN problem with a population size limit of 20,000 was not as intensive as that required for the corresponding ACN problem. The smaller memory usage may be a result of the smaller problem size.

By examining the partitioned problem results by crew type, it is evident that the largest improvements were for the EFN problems with a population size of 20,000 pairings. The partitioned EF problems also had better final solutions than the ACN problems for the cases with population size limits of 2,000 and 4,000 pairings. The original, partitioned EF crew schedules may have been the least optimal and, thus, may have had a larger potential for improvements. The

TABLE 2 Summary of Results for Generated Problems

Number of Duties	Total Cost ($)		Improvement (%)		Calculation Time (s)	
	Heuristic	Optimal	Heuristic	Optimal	Heuristic	Optimal
26	7,074	7,074	17.7	17.7	15	79
28	7,280	7,280	18.4	18.4	22	889
30	7,834	7,792	19.9	20.3	31	6,284

TABLE 3 Summary of Results

Problem	Number of Duties	Population Size, n (pairings)	Improvement (%)	Calculation Time (h)	Memory Usage (MB)
ACN	1,178	2,000	4.5	9.7	61.3
ACN	1,178	4,000	7.4	29.3	148.6
ACN	1,178	20,000	15.8	318.6	341.5
EFN	1,044	2,000	12.0	14.8	30.2
EFN	1,044	4,000	21.0	42.6	67.7
EFN	1,044	20,000	25.7	330.9	210.1
ACS	1,067	2,000	3.0	18.8	79.8
ACS	1,067	4,000	8.2	34.2	151.6
ACS	1,067	20,000	18.0	62.7	306.5
EFS	1,049	2,000	10.2	11.2	33.2
EFS	1,049	4,000	18.3	21.4	54.7
EFS	1,049	20,000	27.9	86.3	516.4
AC	2,245	2,000	2.9	333.6	906.9
EF	2,093	2,000	3.0	266.3	885.4

partitioned EF problems were also smaller than the partitioned AC problems, which may have contributed to better improvements.

As the population size increased, the calculation time and memory usage also increased. However, a direct relationship between the calculation time and memory usage was not observed. Although, in general, problems that required longer calculation times also required more intensive memory usage, these variables were not directly related. The memory usage was found to be related to the number of duties, the population size, and the fitness of the random pairings generated by the heuristic.

The results obtained by solving the full networks confirm that as the number of duties increases, the heuristic becomes less effective at improving the crew schedules. The larger problem, the AC problem, had the smallest improvement in total cost, whereas the EF problem, which was nearly as large as the AC problem, had the second smallest improvement in total cost. These problems also required much longer calculation times, on the order of days, compared with the calculation times required for the smaller problems with the same population sizes. As expected, the resulting memory usage required for such large problems was far more intensive than that required for the smaller, partitioned problems. Increases in the population sizes for these problems were ultimately limited by computer memory constraints.

Sensitivity Analysis

The results of sensitivity analysis presented in Figure 3 and Figure 4 show that as the numbers of pairings per iteration increase, the memory usage and calculation time increase as well. A further increase to the number of pairings per iteration would cause the calculation times to exceed weeks from the start to convergence. Although calculation times may be irrelevant for the practical, long-term application of a final solution, memory usage could be a realistic limitation because stored pairings would fill the random access memory of a typical computer. Any more data that are stored would then need to be accessed from the hard disk, significantly increasing the calculation time for the program.

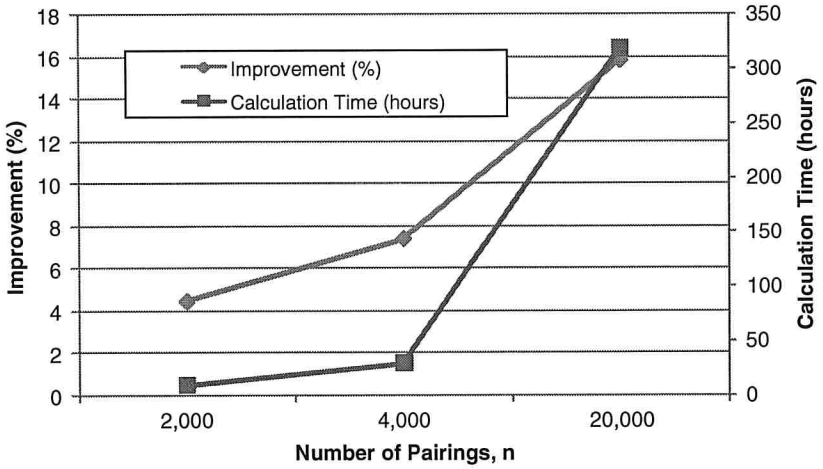


FIGURE 3 Improvement in total cost and calculation time versus number of pairings for ACN problem.

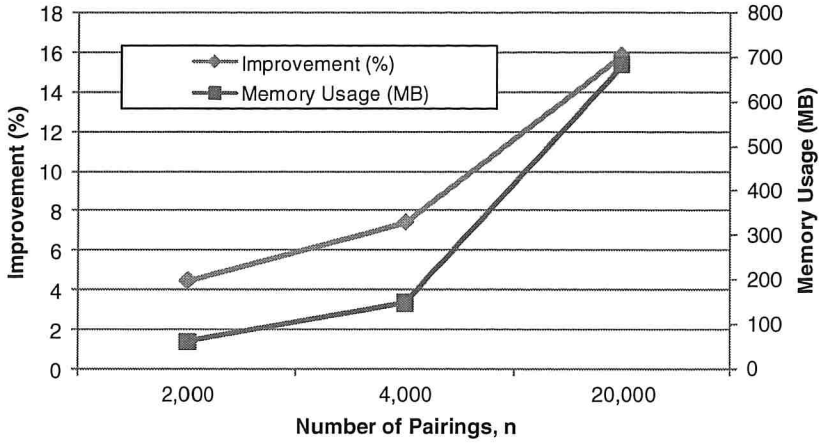


FIGURE 4 Improvement in total cost and memory usage versus number of pairings for ACN problem.

Each of the different data sets behaves somewhat differently as the number of pairings per iteration increases. When the number of pairings is doubled from 2,000 to 4,000 in the ACN problem, the improvement in total cost is moderate. When the pairing limit is further increased by 10-fold to 20,000, the improvement is less. As the improvement increases, the calculation time increases at a faster rate. If calculation time is not of concern, it appears that an appropriate limit on the number of pairings is approximately 20,000.

Improvements in total cost, calculation time, and memory usage vary with the population size, *n*. As the population size limit increases, the CSP becomes more difficult to solve for several reasons. First, increasing the population size makes it more difficult to find feasible pairings. Feasible pairings are those pairings that do not violate any of the problem constraints and that are not already in the restricted subset of pairings. Because many good pairings are likely to have been found and to have entered the restricted subset in later iterations, it becomes increasingly difficult to find new pairings to enter into the subset.

Second, a larger population size means that the restricted subset is also larger. The final IP that must be solved by CPLEX has more pairings, which in turn results in longer calculation times. In all four partitioned problems, the calculation time increased as the population size increased.

Third, because more pairings must be stored and the larger IPs are more difficult to solve, the overall computer program requires more intensive memory usage. Again, the results from all four partitioned problems indicated increasing memory usage as the population size increased.

CONTRIBUTIONS

First, this work contributes to the limited research on passenger rail crew scheduling in North America. As discussed above, the existing literature on passenger rail crew scheduling has primarily been for applications in Europe and Asia, where there are stronger passenger rail markets.

Second, many other rail crew scheduling applications do not include the complex payment regulations that are in the problem used in this work. Few other CSPs have the case in which there is overtime pay, guaranteed salary, away pay, and payments for lodging and meals. The constraints specific to the Amtrak problem are not applicable to many other cases.

Third, this work shows the potential for the use of genetic algorithms to drive a column generation heuristic to solve CSPs. Although genetic algorithms have been used in past CSP solution methods, this work specifically applies a genetic algorithm to the pricing problem to facilitate the column generation process.

CONCLUSION

The existing Northeast Corridor T&E crew schedule was originally developed by the use of commercial crew scheduling optimization software. Once the system is optimized, Amtrak periodically modifies the crew schedule to accommodate various needs, such as changes to the train schedule. The results of this study support the idea that there are potential cost savings in the crew schedule that are no longer optimized. The annual savings for the CSPs defined in this study were found to be significant. These CSPs were formulated from the Amtrak crew schedules but did not consider constraints such as crew seniority or preferences, crew vacation and holidays, and crew comfort. The large potential savings may be a result of the fact that the solutions violated constraints not captured in the study. The actual potential savings are likely to differ.

Since this study was conducted, the Rail Safety Improvement Act of 2008 was signed into law and went into effect in July 2009. The act mandates stricter regulations on crew schedules in an effort to reduce crew fatigue and improve safety. The study results also support the continued use of commercial crew scheduling optimization software to maintain the crew schedule, in light of such changes to crew working constraints.

REFERENCES

1. Bennett, B., and R. Potts. Rotating Rosters for a Transit System. *Transportation Science,* Vol. 2, 1968, pp. 14–33.
2. Wren, A. Computer Scheduling of Public Transport Urban Passenger Vehicle and Crew Scheduling. In *Lecture Notes in Economics and Mathematical Systems,* Vol. 505. Springer-Verlag, Berlin, 1981.
3. Arabeyre, J., J. Fearnley, F. Steiger, and W. Teather. The Airline Crew Scheduling Problem: A Survey. *Transportation Science,* Vol. 3, 1969, pp. 140–163.

4. Bodin, L., B. Golden, A. Assad, and M. Ball. Routing and Scheduling of Vehicles and Crews—The State of the Art. *Computers and Operations Research,* Vol. 10, No. 2, 1983, pp. 63–211.
5. Marsten, R., and F. Shepardson. Exact Solution of Crew Scheduling Problems Using the Set Partitioning Model: Recent Successful Applications. *Networks,* Vol. 11, 1981, pp. 165–177.
6. Smith, B., and A. Wren. A Bus Crew Scheduling System Using a Set Covering Formulation. *Transportation Research, Part A,* Vol. 22A, No. 2, 1988, pp. 97–108.
7. Falkner, J., and D. Ryan. Set Partitioning for Bus Crew Scheduling in Christchurch. *Proc., Computer-Aided Transit Scheduling: Fifth International Workshop,* 1992, pp. 359–378.
8. Caprara, A., M. Fischetti, P. Guida, P. Toth, and D. Vigo. Solution of Large-Scale Railway Crew Planning Problems: The Italian Experience. *Computer-Aided Transit Scheduling,* Vol. 430, 1999, pp. 1–18.
9. Caprara, A., M. Monaci, and P. Toth. A Global Method for Crew Planning in Railway Applications. *Computer-Aided Scheduling of Public Transport,* Vol. 505, 2001, pp. 17–36.
10. Makri, A., and D. Klabjan. A New Pricing Scheme for Airline Crew Scheduling. *INFORMS Journal on Computing,* Vol. 16, No. 1, 2004, pp. 56–67.
11. Desrochers, M., and F. Soumis. A Column Generation Approach to the Urban Transit Crew Scheduling Problem. *Transportation Science,* Vol. 23, No. 1, 1989, pp. 1–13.
12. Lavoie, S., M. Minoux, and E. Odier. A New Approach for Crew Pairing Problems by Column Generation with an Application to Air Transportation. *European Journal of Operational Research,* Vol. 35, No. 1, 1988, pp. 45–58.
13. Barnhart, C., E. Johnson, G. Nemhauser, M. Savelsbergh, and P. Vance. Branch-and-Price: Column Generation for Solving Huge Integer Programs. *Operations Research,* Vol. 46, 1998, pp. 316–329.
14. Fores, S., L. Proll, and A. Wren. A Column Generation Approach to Bus Driver Scheduling. In *Transportation Networks: Recent Methodological Advances,* Pergamon Press, Oxford, United Kingdom, 1998, pp. 195–208.
15. Freling, R., R. Lentink, and A. Wagelmans. A Decision Support System for Crew Planning in Passenger Transportation Using a Flexible Branch-and-Price Algorithm. *Annals of Operations Research,* Vol. 127, 2004, pp. 203–222.
16. Hoffman, K., and M. Padberg. Solving Airline Crew Scheduling Problems by Branch-and-Cut. *Management Science,* Vol. 39, No. 6, 1993, pp. 657–682.
17. Graves, G., R. McBride, I. Gershkoff, D. Anderson, and D. Mahidara. Flight Crew Scheduling. *Management Science,* Vol. 39, No. 6, 1993, pp. 736–745.
18. Ernst, A., H. Jiang, M. Krishnamoorthy, H. Nott, and D. Sier. An Integrated Organization Model for Train Crew Management. *Annals of Operations Research,* Vol. 108, No. 1/4, 2001, pp. 211–224.

The Intercity Passenger Rail Committee peer-reviewed this paper.

Deploying Business Logic for Railway Infrastructure Charging

Rosário Macário, Paulo Teixeira, and Werner Rothengatter

The policy of the European Commission indicates that charges for rail infrastructure should be related to the costs actually incurred, define incentives for more efficient use of the infrastructure, avoid discriminating against users, and allow public authorities to verify compliance with these objectives. European Commission Directives 2001/12/EC and 2001/14/EC define a general framework for the establishment of charges for the rail infrastructure; however, enormous difficulties have arisen in calculating marginal costs and allocating the full costs of the infrastructure to cost drivers. Over the years and as the need for infrastructure managers (IMs) to adopt a more commercially oriented attitude developed, it became vitally important for every IM to understand the link between accounting and charging, requiring the adoption of a business logic that ensures that the cost drivers are well identified and controllable and that investments are made by taking into account future needs. In fact, to face future needs, the cost accounting framework for railway infrastructure should deliver important background information for establishing charges. The framework should also allow comparison of the costs and revenues for each market segment in determining the services on which business activities should be focused and the levels of public funding required to fulfill public service obligation agreements as part of multiyear contracts with predefined levels of service. Finally, cost accounting frameworks for the railway infrastructure should also deliver accurate information for the regulator to ensure equity and fairness in railway market access. The main task for the regulator will be to check that the resulting differentiation of track prices is free of discrimination.

A policy of the European Commission (EC) indicates that charges for the rail infrastructure should be related to the costs actually incurred (*1*), define incentives for the more efficient use of the infrastructure, avoid discriminating against the users, and allow public authorities to verify compliance with these objectives. Directives 2001/12/EC and 2001/14/EC have established a general framework for setting charges for the rail infrastructure; however, enormous difficulties with the calculation of marginal costs and the allocation of the full costs of the infrastructure used to the different cost drivers have arisen. It is vitally important for every infrastructure manager (IM) to understand the link between accounting and charging. This requires the adoption of a business logic that ensures that cost drivers are well identified and controllable. Figure 1 illustrates how important this knowledge is for several modes, given the serious deficits, and in particular, for railways.

In fact, to face future needs, the cost accounting framework for the railway infrastructure should deliver important background information for the establishment of charges. The framework should also allow comparison of the costs and revenues for each market segment as a basis for determining the services on which business activities should be focused and determining the levels of public funding required to fulfill public service obligation (PSO) agreements as part of multiannual contracts with predefined levels of service. Finally, cost accounting frameworks for the railway infrastructure should also deliver accurate information to regulators to ensure equity and fairness in railway market access. The main task for the regulator will be to check that the resulting differentiation of track prices is free of discrimination.

Knowledge of the combination of cost categorization and cost accounting units is essential to understanding how much of each cost category may be tied to incremental charges. This also allows greater flexibility when more complex approaches to railway cost accounting frameworks that may simultaneously fulfill management, charging, and regulation purposes need to be addressed. In brief, there is a need to have a structured cost-accounting scheme that includes capacity costs, that is, life-cycle costs (LCCs) for future maintenance, reinvestment, and investment activities, as well as running costs. This paper presents the final results of the RailCalc project, commissioned by the EC to develop a guide to best practices in complying with EC Directive 2001/14/EC regarding railway infrastructure use charges. RailCalc developed an approach to pricing based on forward-looking incremental costs and proposed the use of this approach to the EC and practitioners. At the time of writing of this paper, the feedback from the practitioners has been positive, and it is possible to understand the acceptance of the principles proposed here in the context of the practices of the European industry. The work also revealed that current European Union (EU) directives do not prevent the industry from adopting and deploying this business logic, as it is not in conflict with EU policy.

TRIGGER FOR A BUSINESS LOGIC

Policy makers increasingly consider railways (once again in transportation history) to be an important part of the transportation system and now also look at the potential of railways to bring increased energy and environmental sustainability to the whole transportation

R. Macário, CESUR, Instituto Superior Técnico, Universidade Técnica de Lisboa, Departamento de Engenharia Civil e Arquitectura (DeCIVIL), Av. Rovisco Pais, Lisbon 1049-001, Portugal, and TIS.PT, Consultores em Transportes, Inovação e Sistemas, s.a., Av. Marquês de Tomar 35, 6°, Lisbon 1050-153, Portugal. P. Teixeira, CESUR, Instituto Superior Técnico, Universidade Técnica de Lisboa, Departamento de Engenharia Civil e Arquitectura (DeCIVIL), Av. Rovisco Pais, Lisbon 1049-001, Portugal. W. Rothengatter, Karlsruher Institut für Technologie, Kaiserstrasse 12, Karlsruhe 76131, Germany. Corresponding author: R. Macário, rosariomacario@civil.ist.utl.pt.

Transportation Research Record: Journal of the Transportation Research Board, No. 2159, Transportation Research Board of the National Academies, Washington, D.C., 2010, pp. 44–51.
DOI: 10.3141/2159-06

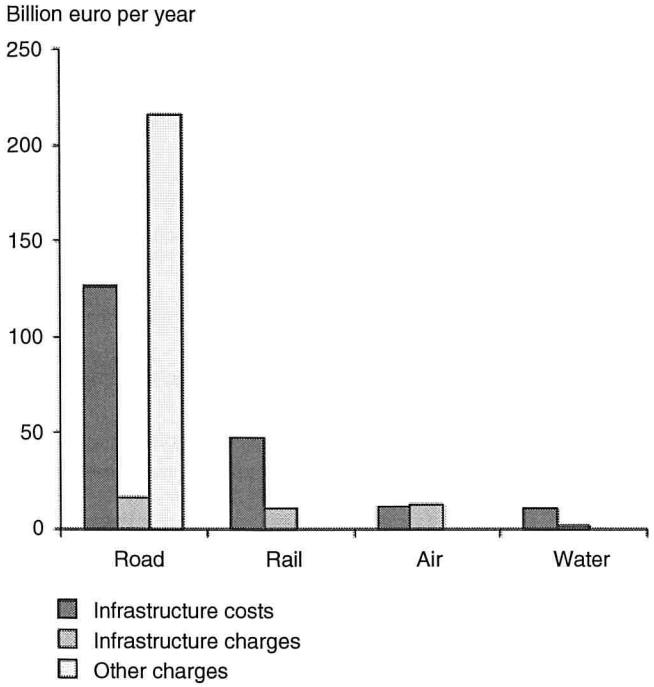


FIGURE 1 Annual infrastructure costs and charges (billions of euros, 2005) (*2*).

system. This is reflected in the current concerns of European transportation and energy policies, which devote more attention to the role of railways, considered a promising alternative to congested and polluting road traffic, and which address components of the railway system such as freight transportation, which has been performed by road haulage for decades.

Some of the key challenges for the railway sector lie in improving its performance in such areas as availability, reliability, cost-effectiveness, safety, and comfort (*2*). In support of these goals, the European Railway Directives Package promotes vertical separation, along with prospects for the future privatization of at least parts of the system (*1*).

This should allow responsibilities to be distinguished more clearly, fostering competition in the rail transportation market. To this extent, the European Commission has proposed a series of railway restructuring measures (known as the "infrastructure package") that look at separating train operations from infrastructure management. From a cost management perspective, the challenges for IMs are growing, as there is an increasing demand for gains in efficiency. These have implications for both operational cost-efficiency and decisions related to capital investments. Table 1 presents a snapshot of the current charging situation for railways in Europe, as surveyed in the RailCalc project, which provides evidence of the diversity of charging practices and which prevents the rigorous comparison of performance.

Historically, capital costs have been supported by public funds. Although that situation is not bound to change significantly in the near future, several factors support the view that public funding should be reduced, insofar as cost optimization allows it and to the extent that the railway market can bear to pay above-marginal costs for specific segments as part of a lean cost management system that is adapted to the business of IMs as well. Because public funding for railways has increasingly been subject to discussion, market-based revenues may play a predominant role in the future of the railway sector. Indeed, the EU railway sector is currently subsidized to the sum of €73 billion per year (*2*).

According to recent data, the financial benefit to the sector is split almost equally between subsidies for the infrastructure and subsidies for fare reductions but with infrastructure subsidies constituting the largest share (€37 billion) (*2*). It is interesting to observe the degree of such dependence by evaluation of the cross-sector situation. From Figure 1 it is clear how differently each sector performs when it comes to covering infrastructure costs with charges (either direct or indirect).

Therefore, in comparison with other sectors, it is evident that the rail infrastructure has difficulties in reaching financial sustainability through charges, set in such a way that cost-efficiency objectives are not counteracted. This means that it is increasingly important that IMs operate on the basis of the logic used by the private sector, even if the company is owned by the state. This requires investment in projects in which the highest return on the capital invested can be expected. If network extensions and upgrades or a guarantee of regional access are, in the first instance, motivated by public interest, funding by public authorities will be necessary. In the future, this funding should be subject to contractual arrangements between the IM and the state that effectively promote cost-efficiency.

Therefore, in the long run, all costs must be covered, one way or the other. Costs can be recovered either through PSO agreements linked to contract programs established with the state or by revenues from track access charges and other services provided by the IM. Given its natural monopolistic character, because of the high cost associated with infrastructure, there is also a need for regulation that cannot be detached from the transparency that cost accounting systems are ever more required to feature. It is, however, worth indicating that despite being an industry prone to the development of natural monopolies, in particular, in the provision of infrastructure, an effective competition between railways and the road sector throughout the world maintains some competitive pressure on the railway companies. Tables 1 to 3 give an overview of the state of the art for all European countries surveyed as part of the RailCalc project.

COST ACCOUNTING FRAMEWORK

It is vitally important for every IM to understand the link between accounting and charging, which requires the adoption of a business logic that ensures that the cost drivers are well identified and controllable.

To face future needs, the railway infrastructure cost accounting framework should not only deliver important background information for the establishment of charges but also allow comparison of the cost and revenues for each market segment as a basis for determining the services on which business activities should be focused. The levels of public funding required to fulfill PSO agreements as part of multiannual contracts with predefined levels of service also need to be known. Finally, railway infrastructure cost accounting frameworks should also deliver accurate information to the regulator to ensure equity and fairness in railway market access. The main task for the regulator will be to check that the resulting differentiation of track prices is free of discrimination.

To this extent, the review of current accounting practices allowed the identification of best practices for cost categorization and the adoption of cost centers deemed compatible with the concerns mentioned above. For cost categorization, it is desirable to maintain a clear distinction between infrastructure-related cost categories

TABLE 1 Current Practice in EU

Country	Pricing Principle	Charge Calculation Procedure	Main Variables Considered (other than traffic output units)
Austria	MC+	Average marginal costs (econometric) complemented with markups	Infrastructure, railway service, rolling stock, traffic, time
Belgium	FC–	Allocation of a predefined set of costs (cost target) to traffic output according to demand- and supply-related variables	Infrastructure, railway service, rolling stock, operation, traffic, time
Bulgaria	MC	Allocation of marginal costs (estimation) to traffic output according to supply- and demand-related variables	Infrastructure, railway service
Czech Republic	MC+	Allocation of marginal costs (estimation) to traffic output according to supply- and demand-related variables	Infrastructure, railway service
Denmark	MC+	Average marginal costs (estimation) complemented with markups	Infrastructure, railway service, time
Estonia	FC	Allocation of total costs to traffic output	Railway service
Finland	MC+	Average marginal costs (econometric) complemented with markups	Railway service, rolling stock
France	MC+	Allocation of a predefined set of costs (including MC and markups) to traffic output according to supply- and demand-related variables.	Infrastructure, railway service, traffic, time
Germany	FC–	Allocation of a predefined set of costs (cost target) to traffic output according to supply- and demand-related variables	Infrastructure, railway service, rolling stock, operation, traffic
Greece	MC	Allocation of marginal costs (estimation) to traffic output according to supply-related variables	Infrastructure, rolling stock, operation, time
Hungary	FC–	Allocation of a predefined set of costs (cost target) to traffic output according to supply- and demand-related variables	Infrastructure, railway service
Italy	FC–	Allocation of a predefined set of costs (cost target) to traffic output according to supply-related variables	Infrastructure, rolling stock, operation, time
Latvia	FC	Allocation of total costs to traffic output	Railway service
Lithuania	FC	Allocation of total costs to traffic output	Infrastructure
Luxembourg	MC	Allocation of marginal costs (estimation) according to demand-related variables	Railway service, rolling stock
Netherlands	MC	Allocation of marginal costs (estimation) to traffic output	—
Poland	FC–	Allocation of a predefined set of costs (cost target) to traffic output according to supply- and demand-related variables	Infrastructure, railway service, rolling stock, operation
Portugal	MC+	Allocation of marginal costs (estimation) to traffic output according to supply- and demand-related variables	Infrastructure, railway service, operation
Romania	FC–	Allocation of a predefined set of costs (cost target) to traffic output according to demand-related variables	Railway service
Slovakia	FC–	Allocation of a predefined set of costs (cost target) to traffic output according to supply- and demand-related variables	Infrastructure, railway service
Slovenia	FC–	Allocation of a predefined set of costs (cost target) to traffic output according to supply- and demand-related variables	Infrastructure, rolling stock, operation
Spain	MC+	Allocation of marginal costs (estimation) to traffic output according to supply- and demand-related variables	Infrastructure, rolling stock, traffic, time
Sweden	MC+	Average marginal costs (econometric) complemented with markups	Railway service, rolling stock
Switzerland	MC+	Average marginal costs (estimation) complemented with markups	Railway service, time
United Kingdom	MC/MC+	Allocation of marginal costs (estimation) to traffic output according to supply- and demand-related variables	Railway service, rolling stock, operation, traffic, time

NOTE: MC = marginal cost; FC = full cost; — = dimensionless.

supported by common definitions regarding cost item depreciation, upgrades, renewals, maintenance, and management and operations. Such a detailed categorization should be combined with the adoption of accounting cost centers that act as building blocks defined at the bottom levels of the physical and organizational infrastructure setting. This means a definition at the line section level for track-related cost drivers (e.g., maintenance and renewal associated with traffic) and at the organizational unit level for service-related cost drivers (e.g., management and operations).

This combination of cost categorization and cost accounting units is essential to understanding how much of each cost category may be tied to incremental charges. This also allows greater flexibility in addressing more complex approaches to railway cost accounting frameworks that may simultaneously fulfill management, charging, and regulatory purposes. In brief, a structured cost accounting scheme that includes capacity costs (LCCs for future maintenance, reinvestment, and investment activities), as well as running costs, is needed.

This recommendation on best practices for the cost structure aims to bring a greater business focus to cost accounting, which enables better insight into the costs of activities entailed in the provision of a network to be gained (*3*). However, conventional cost accounting

TABLE 2 Cost of Use of Asset-Related Charges

Country	Charge	Component	Units	Variable Name
Austria	Quality and line-related markups and markdowns	Wear on tracks by traction units affecting charges	MU/train km	Traction unit category
Austria	Usage charge	Line-related usage charges	MU/train km	Line category
Austria	Usage charge	Charge in the amount of directly incurred costs	MU/gross tonne km	
Belgium	Train path-line charge	Coefficient of mass on rail admissible for train path	—	Weight of train Section operational importance Section reference speed
Bulgaria	Railway infrastructure use	Use of permanent way	MU/gross tonne km	Type of train services Line category
Czech Republic	Use of intrastate, nationwide, and regional railway infrastructure	Price for ensuring infrastructure operability	MU/1,000 gross tonne km	Type of infrastructure Type of train services
Denmark	Infrastructure charges	Kilometer charge	MU/train km	
Estonia	Track access charge	Variable cost-related fee	MU/gross tonne km	
Finland	Infrastructure charge	Basic charge	MU/gross tonne km	Type of train services
France	Minimum services	Running charge	MU/train km	Type of train services
Germany	Infrastructure use (1 year)	Track charge	MU/m of track year	
Germany	Infrastructure use (1 year)	Catenary charge	MU/m of catenary year	
Germany	Infrastructure use (1 year)	Switch charge	MU/switch year	Switch type, switch category
Germany	Train path	Basic price	MU/train km	Route category Gross weight of hauled load
Greece	Basic fee for use of infrastructure	Base price corresponding to line's maintenance	MU/train km	Line section Vehicle speed range Axial load range Number of axles
Hungary	Basic services	Charge for running trains	MU/train km	Type of train services Line category
Italy	Access charge	Use of infrastructure	MU/train km, MU/min	Usage parameter (speed, weight, number of pantographs)
Latvia	Charge for use of infrastructure	Use of railway infrastructure for one train kilometer	MU/train km	Type of train services
Lithuania	Train traffic	Unit rate	MU/train km	
Luxembourg	Minimal services	Directly ascribable cost with exploitation	MU/train km	Type of train services Train category Number of coaches Mass category
Netherlands	Basic access package	Tariff per train kilometer	MU/train km	
Netherlands	Basic access package	Tariff per tonne kilometer	MU/tonne km	
Poland	Basic charge	Unit rate assigned to individual sections of railway line	MU/train km	Maximum speed of the section Type of train Average technical vehicle speed Gross weight
Portugal	Essential train services	Basic tariffs	MU/train km	Type of traffic Line Type of route
Romania	Infrastructure access charge	Minimum service package	MU/train km	Type of service
Slovak Republic	Maximum charge for access to infrastructure	Maximum charge for access to infrastructure	MU/train km	Track category type of traffic
Slovak Republic	Maximum charge for thousand gross tonne kilometers	Maximum charge for thousand gross tonne kilometers	MU/1,000 gross tonne km	Track category type of traffic
Slovenia	User charge	Coefficient for track wearout	—	Type of train Type of line
Spain	Running	Running charge	MU/train km	Type of line Type of service
Sweden	Marginal-cost-based charges incurred by operation of traffic	Track charge	MU/gross tonne km	
Switzerland	Basic service	Minimum price (maintenance)	MU/gross tonne km	Type of train services
United Kingdom	Variable usage charge	Variable usage charge	MU/train mile	Vehicle class

NOTE: MU = monetary unit; — = dimensionless factor.
SOURCE: IM network statements.

TABLE 3 Use of Markups

Country	Charge	Component	Units
Austria	Quality and line-related markups and markdowns	Wear on tracks by traction units affecting charges	MU/train km
Denmark	Bridge tariff	Infrastructure charges	MU/train
Estonia	Traffic charge	Traffic charge	MU/100 seats km
Estonia	Capacity reservation charge	Capacity reservation charge	MU/path km
Finland	Infrastructure charge	Investment tax	MU/gross tonne km
Finland	Infrastructure charge	Infrastructure tax	MU/gross tonne km
France	Reservation charges	Path reservation charge	MU/path km month
Sweden	Train path charge	Special charge for passenger traffic	MU/gross tonne km
Sweden	Train path charge	Special charge for Oresund link	MU/train
Switzerland	Contribution margin	Contribution margin of franchise holders' trains	% revenue
Switzerland	Contribution margin	Contribution margin of nonfranchise holders' trains	MU/train km
Switzerland	Contribution margin	Other services charged via contribution margin	MU/train
United Kingdom	Fixed track access charge	Fixed track access charge	MU

systems for railways have not focused on the delivery of such detailed information. Therefore, important improvements to state-of-the-art cost management systems and information systems in association with best practices on cost structuring need to be considered.

Other challenges to cost accounting that have a direct influence on the implementation of recommended best practices were identified. Some of these challenges include the asset valuation methodologies used, the means of estimating running costs, and the means of allocating costs, including the different costing approaches typically considered (*4*).

One of the most important challenges to cost accounting for charging in the spirit of Directive 2001/14/EC has been the development of a proper means of calculation of the differentiated costs incurred during the provision of a network in support of the charging decisions, often interpreted as the short-run marginal cost (SRMC) level. It is therefore necessary to understand the costs incurred by the activities of the IMs, their nature, and the corresponding drivers. This is an area of research clearly associated with two main different approaches: the econometric approach (top–down costing) and the engineering approach (bottom–up costing).

Although marginal costs for the provision of a network may play a role, it is of utmost importance to understand the need to change the definition as part of the change in the decision-making unit of the IM. It cannot continue to be the train, car, or axle kilometer; instead, it should be the group of slots that must be made available for the successful provision of a particular railway service.

Nevertheless, previous and current research shows that it is difficult to define incremental costs for each cost category (*5*). It is thus necessary to simplify and homogenize calculation procedures by the use of incremental cost formulations that may provide a better understanding of the cost drivers. This will encourage cost-efficiency and improve the IM's financial performance. Therefore, innovative cost management approaches are required for two major reasons:

- Internally, to cut unnecessary activity costs, which will improve the cost-efficiency of the IM through better cost management and reaching higher-value market segments, and
- Externally, to identify sound reasons (such as enhanced economic sustainability, in which the environment and energy play important roles) for justifying public funding and subsidies for the remaining costs not covered by charges set in such a way that cost-efficiency goals are not counteracted. To accomplish this objective, it seems to be indispensable to know how close a service comes to breaking even by considering the market revenues and the state service contributions on the revenue side.

Given this change context, RailCalc concluded that it has become crucial to link cost accounting to the whole management approach to infrastructure management. It is necessary to move from the use of conventional cost accounting systems to the use of new approaches focused on delivering the sort of detail and deeper knowledge of cost on which businesses depend, enabling effective benchmarking.

Improving conventional cost accounting practices on the basis of better knowledge of usage-related costs and their nature brings to the discussion the application of activity-based costing (ABC) by railway IMs.

ABC represents a cost accounting system based on the current costs of activities rather than historic costs, and its use is therefore advisable to reconcile top–down and bottom–up costing approaches. Moreover, ABC may take full advantage of the sophistication effort on bottom–up economic and engineering models, which in all cases are required to feed the ABC cost-partitioning mechanisms.

The application of such a cost accounting and allocation concept may help clarify the somewhat blurred area that lies between shorter-term and longer-term cost assessments by providing an approach to pricing based on forward-looking incremental costs. From a managerial perspective, ABC also provides the ground for activity-based management (ABM). ABM focuses on managing activities to reduce costs and improve customer value and can be divided into operational and strategic ABM (*6*), as follows:

- Operational ABM (i.e., doing things right) uses ABC information to improve efficiency. Those activities that add value to the IM product can be identified and improved. Activities that do not add value need to be reduced to cut costs without reducing product value.
- Strategic ABM (i.e., doing the right things) uses ABC information to decide which products should be developed and which activities should be used. In the case of railway IMs, this also means anticipating market trends, which enables planning for new investments and provides decision-making support for long-term PSO contracts and multiannual contracts, ensuring that long-term deficits are supported by public funding.

Such an approach to the structuring of cost accounting can contribute to ensuring that the expenditures of IMs are allocated to the marketable services of the IMs, thus ensuring the cost relatedness of charges. Combined with proper cost allocation partition keys (based on marginal costing approaches to network usage parameters), ABC may provide a pragmatic approach to justifying charges on an incremental cost basis. Moreover, it may allow more frequent and effective cost benchmarking, especially for the allocation of common costs, which would allow regular cross comparisons of costs among IMs and which would fulfill the information requirements for railway regulators to accomplish their mission of surveillance for nondiscriminatory practices.

CHARGING FRAMEWORK

The application of SRMC approaches faces obstacles when it comes to practical implementation, as it is difficult to establish criteria related to how much of each cost category may be deemed variable for the calculation of marginal costs. However, charging according to incremental costs, understood as SRMC, implies that the IM is unable to raise the necessary funds to develop its activity, requiring contributions from the state budget to finance railway network developments and improvements. Moreover, it does not generate a clear link to improved IM cost-efficiency and cost-optimization strategies. Taking these concerns into consideration, the RailCalc project reference framework for charging has assumed that railway IMs need to adopt a business-oriented approach toward the future. Therefore, it is increasingly important that the charging system adopt a dynamic forward-looking approach to current and future costs rather than an approach based on historic costs. Current practices were thus assessed on the basis of this reference framework.

Indeed, despite developments in the calculation of marginal costs that have been achieved in some research studies, IMs often consider incremental costing to be complex (*5, 7*), although a significant diversity of practices and interpretations of Directive 2001/14/EC exists. On the other hand, current practices also bring about potential drawbacks, namely, the risk of uncontrolled overpricing, in particular, if an incumbent detains infrastructure facilities also used by competitors, as illustrated in Figure 2.

Although Directive 2001/14/EC indicates that the application of cumulative charges cannot exceed total costs (SRMCs plus fixed costs), uncontrolled overpricing may arise. The result of this will be

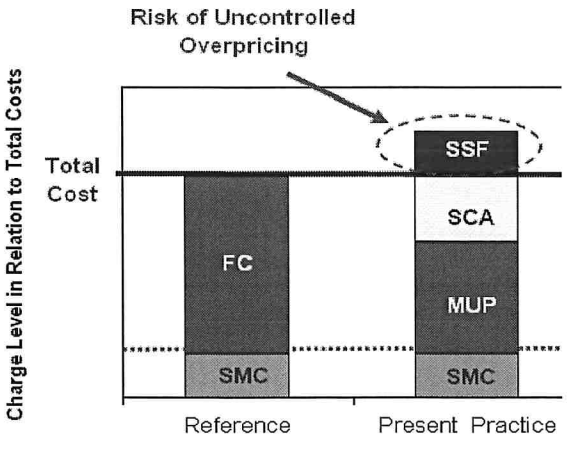


FIGURE 2 Charge in relation to total cost (*B*) (FC = full cost; SSF = special service fee).

the distortion of overall charging when these added charges for special services are set as an obstacle to new entrants, especially when scarcity charges and markups are applied and already cover total costs.

Consequently, the exclusion of competitors could also take place, even in the case of supposedly fair charging for all infrastructure components,

- If the essential facility is not open for use in the case of capacity constraints (e.g., the incumbent uses all of the possible slots on a part of the infrastructure and is not willing to share the paths because of grandfather rights),
- In the case of time constraints (e.g., all economically attractive slots within a certain period of time are allocated to restricted parties), and
- Under technical constraints (e.g., the infrastructure can be used only by a certain kind of train with a certain technical design, without any interoperability with other parts of the network).

These examples emphasize the need for the regulator to control access not only in terms of primary charges (minimum access package) but also on the basis of additional charges.

Another drawback is the fact that the adoption of scarcity charges (SCA) and markups (MUP) is associated with the willingness to pay (WTP) of train operating companies. Moreover, current charging practices allow the addition of WTP-related elements to basic social marginal cost (SMC) charges. This means that WTP prevails over SMC considerations.

Therefore, whenever WTP considerations are part of charging practices (no matter if they are through markups or scarcity charges), marginal cost considerations and calculations become useful merely as a general indication for charging. The overall charge in this case is actually determined by the acceptance by the train operating company (TOC) of a certain charge bearable by the market, which overrides SMC considerations.

This dynamic is illustrated in Figure 3, which depicts the contradictory effects associated with current practices in relation to the cumulative use of different charging elements. Tables 4 and 5 provide evidence of these practices.

In this context, the adoption of ABC systems in railways, according to the RailCalc proposition illustrated in Figure 3, may be considered a pragmatic approach to setting up a basis for a forward-looking incremental cost-based charging system. This proposition keeps the essential spirit behind the existing legal framework, namely, basing the charging principle on incremental costs (Directive 2001/14/EC), although it also has the potential to deliver a better focus on the business orientation of the IM, without compromising fair access conditions.

Infrastructure charging centered on ABC would comprise the marginal operational costs attributable to the level of activity of the IM. Therefore, on top of the charges based on marginal external costs (MECs), added charges should also be set by consideration of marginal operational costs. The use of such a charging approach will depend on market acceptability, that is, WTP considerations, which will define a minimum setting at an indicative SMC reference level.

This proposition represents an improvement over current charging practices, as it develops around the concept of operational cost optimization complemented by WTP as a driver for the application of markups above the SMC level. It also requires the slot charge, including the use of facilities, to be addressed as a whole, depending on market conditions.

The approach might be extended to the point of covering total costs, with longer-term charging remaining within the long-run marginal

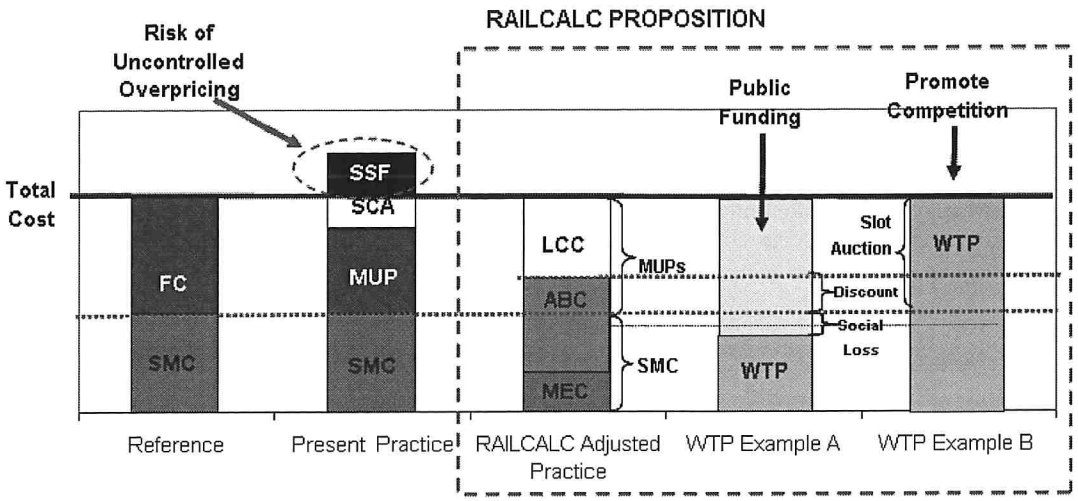


FIGURE 3 Business logic proposition advanced in RAILCALC project (*8*).

cost (LRMC) and LCC concepts. In this circumstance, markups would represent a combination of SRMC and medium-run marginal cost and LRMC concerns, promoting coverage of the cost components, with an impact on the long-term sustainability of the railway activity, namely, ABC-related costs (SRMCs and medium-run marginal costs) and LCC-related costs.

It should be noted that charging below the ABC level and higher than the estimated SMC level is still possible, depending on IM management decisions (e.g., through discounts, in light of the principles of the EC charging directive). This approach is also compatible with the legal framework, to the extent that WTP can be seen to be a compounded measure for slot reservation, SCA, and MUP, leaving the chance for the adoption of discounts and performance schemes.

For the main regulation-related implementation issues, it is important that the market value of the slots offered by the IM be captured above the point at which charges are at the ABC level, which will require the adoption of auctioning mechanisms. From the regulatory point of view, charges above the total cost would further require the establishment of slot reservations for new entrants and could be supported, thus avoiding the prevalence of incumbent operators under exceptional market conditions.

By following the approach already adopted for similar circumstances, namely, in the road sector, on the side of charging regulations, it is suggested that peak pricing in relation to off-peak pricing be capped by a factor of 2. By adopting such a price cap principle, the regulator will ensure that there is an incentive to invest in that expansion of capacity, thus avoiding situations in which the IM could

TABLE 4 Congestion-Related Charging Elements

Country	Charge	Component	Unit
Belgium	Train path line charge	Coefficient of deviation compared to standard train path Coefficient relating to time slot, day, and direction of movement	— —
Greece	Basic fee	Traffic management	—
Italy	Access charge	Discount Use of infrastructure	MU/train km MU/train km
United Kingdom	Capacity charge	Capacity charge	MU/train mile

NOTE: MU = monetary unit; — = dimensionless factor.
SOURCE: IM Network Statements.

TABLE 5 Scarcity-Related Charging Elements

Country	Charge	Component	Unit
Austria	Standard package train movement charge	Markups for time and local capacity bottlenecks	MU/train km
Denmark	Infrastructure charges	Capacity charge	MU/train
Germany	Train path charge	Incentives to raise efficiency	—
Luxembourg	Minimal services	Scarcity of capacity charge (infrastructure congestion)	MU/train km

NOTE: MU = monetary unit; — = dimensionless factor.
SOURCE: IM Network Statements.

simply exploit peak charges in relation to off-peak charges under capacity constraints. Therefore, the thresholds that deserve the attention of regulators are the ABC charge (with evidence required from ABC information systems) and total cost levels.

CONCLUSIONS

An accounting framework based on a systematic ABC rationale would allow the identification of the functional costs that enable charging at a level equal to or above MECs and below the total costs with greater rigor, which would clarify the drivers of each cost category. This constitutes an approach to pricing based on forward-looking incremental costs supported by current costs rather than historic costs. Hence, the application of ABC principles to the IM business may provide multiple contributions to the development of the railway business, namely,

- By providing the level of transparency needed for effective regulatory monitoring (regulatory perspective),
- By enhancing the accuracy of cost accounting and enabling cost management (management perspective),
- By providing an effective link between the cost of activities and charging to send sound economic signals to TOCs (market perspective), and
- By fostering cost-efficiency and reliable market reactions to decisions on investments (investment perspective).

Finally, the current legal framework offers no barriers to the implementation of what was defined in the RailCalc project as a best practice on bridging cost accounting and charging for railway infrastructure services.

REFERENCES

1. Accounting Separation and Cost Accounting in a Liberalised Telecommunications Market. Commission Recommendation 98/322/EC. European Commission, 1998.
2. *Size, Structure, and Distribution of Transport Subsidies in Europe.* Technical Report 3/2007. European Environment Agency, Luxembourg, 2007.
3. *Calculation Transport Infrastructure Costs.* Final Report of the Expert Advisors to the High Level Group on Infrastructure Charging (Working Group 1). European Commission, Brussels, Belgium, April 28, 1999.
4. *Switching Track to IFRS—Embedding International Financial Reporting Standards in the Railway.* PriceWaterhouse Coopers, 2005. www.pwc.com/transport.
5. Generalisation of Research on Accounts and Cost Estimation (GRACE). Deliverable 5: Monitoring Pricing Policy Using Transport Accounts, Version 1.0. GRACE Consortium, June 2007.
6. Kaplan, R. S., and R. Cooper. Cost and Effect: Using Integrated Cost Systems to Drive Profitability and Performance. Harvard Business School Press, Boston, Mass., 1998.
7. Tanczos, K., F. Meszaros, and A. Torok. *Rail Infrastructure Costs in Hungary.* Department of Transport Economics, Budapest University of Technology and Economics, Budapest, Hungary, 2007.
8. RailCalc Project. http://www.railcalc.org. Accessed March 24, 2010.

The Freight Transportation Economics and Regulation Committee peer-reviewed this paper.

New Approach to Appraisal of Rail Freight Projects in South Korea

Using the Value of Freight Transit Time Savings

Kyungwoo Kang, Anne Strauss-Wieder, and Jin Ki Eom

This paper contributes to the estimation of the value of rail freight transit time savings, reliability, and service frequency for producers and transport operators by using stated preference data for short-haul rail freight shipments in South Korea. To appraise investments in freight projects in South Korea, the mixed logit model was used to determine the value of rail freight transit time savings, resulting in an estimated savings of about $2.31/h per shipment. The value of rail freight reliability (the percentage of on-time arrivals and departures) is $2.78/h of delay and $0.54/day for scheduled freight service delay. The results for freight time value for different economic agents are surprising: shippers perceived the value of transit time, transit time reliability, and service frequency to be far higher than producers did. Shippers valued rail freight transit time savings at about $2.35/metric ton h, whereas this same value for producers was $1.21/metric ton h. When the time benefit based on these values was added to the preliminary feasibility evaluation for 20 freight railroad infrastructure projects in South Korea, the number of projects for which benefits outweighed costs increased from four to 11.

The total route length of roads in South Korea has been growing rapidly at a rate of 2,000 km/year over the last 30 years, resulting in an annual growth rate of 4.2%. However, the route length of the railroad network in South Korea stayed fairly constant over the same period until 2004, when the South Korean high-speed rail network, the South Korea Train Express (KTX), began operations.

The number of kilometers traveled by passengers has been increasing substantially by about 2.8% per year by rail and 4.5% per year by road over the past 30 years. Rail's share of passenger transportation stayed fairly constant until 2004, at which point there was a rapid increase because of the start of high-speed rail operations. The modal split of freight transportation between road and railroad has changed substantially, with the share of freight tons carried on the road increasing from about 60% in 1980 to almost 77% in 2007, whereas rail's share of freight carried decreased from 28% in 1980 to only 6% in 2007.

The reasons for rail's declining share of freight transportation include supply chain management, which requires the timely delivery of freight, and geographical considerations, including the relatively short distances across South Korea, which makes road travel attractive. Another significant reason for the decline of rail's share of freight is capacity constraints for railroad operations, as railroads in South Korea are operating close to their maximum capacity. To increase railroad capacity, the physical capacity must be increased. This can be done by expanding the railroad network by adding new routes and additional tracks and extending electrified lines. From 1997 to 2007, of the $169 billion of public investment in the transportation infrastructure, about 50% was spent on roads, whereas only 15% was spent on rail, including high-speed passenger rail projects.

The practice of carrying out preliminary feasibility studies was introduced in 1999 to encourage the use of a cautious approach to new large-scale projects (those costing over about $50 million) meant to enhance the efficiency of public investment by verifying the feasibility of projects from the viewpoints of economic feasibility, policy objective, investment priorities, timing, and financing methods. One basis for such preliminary feasibility studies was general research on large-scale development projects. Article 38 of the National Finance Act of South Korea provides the legal basis for each preliminary feasibility study and stipulates that projects requiring not only the approval of the Ministry of Planning and Budget but also that of the National Assembly must undergo inspection and verification to corroborate the results of the preliminary feasibility study.

Unfortunately, most preliminary feasibility studies of rail freight projects required by South Korean law have not accurately assessed economic feasibility, as their cost–benefit analyses (CBAs) are flawed. This is also true elsewhere in the world. Except in rare cases, CBAs do not take into account the time value of rail freight and the resulting benefits of better transit time, its reliability, and service frequency. The main reasons are the lack of a theoretical background and limited data. For example, the government of the United Kingdom does not endorse the use of such freight time benefits and is specifically opposed to consideration of the value of freight time savings for transportation project appraisal; instead, it simply considers traditional operating cost savings (*1*). As the physical distribution system has taken shape and with the development of just-in-time supply chain management and intermodalism, research has begun to focus on the value of freight time as an important benefit. Studies have been confined, however, to European countries and Australia.

The CBAs carried out by the World Bank and Eurobank consider only the operating cost of freight vehicles and the cost of labor. Sweden includes the value of freight time as a benefit that can help justify projects of the National Road Administration and the National Rail Administration. In the case of roads, freight time cost is considered in the CBA. In the case of rail, the value of service reliability is included and applied. In these schemes, freight is classified along simple lines.

K. Kang, Department of Transportation Engineering, Hanyang University, Sa-3-dong, Ansan 425-791, South Korea. A. Strauss-Wieder, A. Strauss-Wieder, Inc., 330 South Chestnut Street, Westfield, NJ 07079. J. K. Eom, Railway Transport and Logistics Research Department, South Korea Railroad Research Institute, Uiwang, Kyonggi 437-757, South Korea. Corresponding author: K. Kang, kyungwoo@hanyang.ac.kr.

Transportation Research Record: Journal of the Transportation Research Board, No. 2159, Transportation Research Board of the National Academies, Washington, D.C., 2010, pp. 52–58.
DOI: 10.3141/2159-07

For example, classifications include bulk versus general cargo, low-density versus high-density goods, and low-value versus high-value goods. Although many CBAs calculate a value for rail freight time savings, such considerations are kept out of CBAs in practice (*2*).

The value of time in the passenger transportation sector has been well studied and extensively applied to maximize a utility function, in the case of freight transportation; on the other hand, the approach has generally been to maximize a profit function. It is, however, difficult to identify the economic agent whose profit function must be maximized. The willingness to pay for freight time savings differs among economic agents, such as consumers, producers, and transport operators. A final difficulty is that most previous studies of the value of freight transit time savings are restricted to roads (41 studies), whereas only five studies have considered the rail sector.

As emphasized by Zamparini and Reggiani, research on freight transportation and the value of time savings should consider the role of additional variables, such as the type of goods transported, the length of the trip, and differences between countries (*3*). To the best of the authors' knowledge, the study described here is the first to evaluate the value of rail freight time savings in an Asian country. This research is also unique compared with previous studies in several respects, such as its distinctions between economic agents (producers versus freight operators), commodities (container versus bulk), and shipment directions (inbound versus outbound). The value of different characteristics of rail freight service, such as transit time, service reliability, and service frequency, is also estimated.

The paper is structured as follows. The next section summarizes previous studies on the value of freight time. The main theoretical and modeling features are then presented. An example is provided, and an empirical application for rail freight operations in South Korea is carried out. Finally, the results are presented and the conclusions are summarized.

PREVIOUS STUDIES ON VALUE OF FREIGHT TIME

Fowkes stated that freight decision-making processes and freight movement contracts are generally kept confidential. This makes it nearly impossible to study the value of freight time by use of the revealed preference (RP) method. Similarly, conventional stated preference (SP) methods are of limited usefulness because the interviewees are usually high-level decision makers in companies (*4*).

Transek gathered SP data about export products transported by trucks in Sweden in 1991 (*5*). By using variables such as door-to-door transportation time, transportation cost, the frequency of late shipments, and loss risk (the frequency of shipping loss), the value of freight time was estimated to be 30 SEK per truck hour ($5.34 in 1991 dollars). Fridstrøm and Madslien estimated that the value of freight time ranged from 0 to 75 NOK ($12.38 in 1992 dollars) per shipment hour by using SP data on different kinds of products handled at a wholesale company in Sweden (*6*). Wynter stated that the use of SP data leads to more accurate estimates of the value of freight time than the use of RP data does (*7, 8*).

Jeong analyzed the effect of economic development on transportation investment (*9*). Jeong used SP data to analyze the effect of road investments on productivity, that is, on manufacturers' distribution costs (*9*). To estimate the value of time in freight transportation in Finland, Kurri et al. collected SP data on road users in 1997 and railroad users in 1998 and stated that the different values in their study were typical (*10*). They estimated the average value of a road freight load [$1.53 per metric ton hour (ton h)], the average value of a rail freight load ($0.10/ton h), the delay value of a road freight load ($47.47/ton h), and the delay value of a rail freight load ($0.50/ton h), with the variables being total transportation time (door to door); total cost, including the cost of transit time and delays; and average unexpected delays. Kawamura (*11*), Wigan et al. (*12*), and De Jong et al. (*13*) also used SP data to estimate the value of freight time.

To estimate the value of time in long-distance freight transportation, Hodkins and Starkie gathered RP data on variables such as prices and transportation time for various goods carried within Australia for long distances in the 1970s (*14*). The estimated value of freight time was AUS$10/ton d ($7.81 in 1988 dollars). Blauwens and van de Voorde cautioned that real and intended behavior must be differentiated when SP data are used (*15*). In general, the real and intended behaviors in freight operations are much more different from the behaviors on the passenger side because the decision-making processes for freight operations are complex. However, it is difficult to make this distinction without additional data. The most recent research used a combined model of integrated SP and RP data. That model explained the data better than the use of SP and RP data on their own did, as shown by Viera (*16*) and Swait et al. (*17*).

Kang used RP and SP data to develop the combined model and used a multinomial probit model to estimate choice behavior (*18*). That research was based on the observation that the qualitative improvement of transportation services takes the form of benefits in a variety of areas, including shipment reliability, transit time, risk of loss or damage, and convenience of use. Under the assumption that the quantity of all freight and the decision-making behaviors of the owners of goods do not change, that research estimated that the increase in the willingness of goods owners to pay for qualitative improvements to transportation services is a measure of the value of freight time. Bruzelius reviewed the research on the valuation of freight time and other quality factors affecting the transportation of goods, as well as how logistics improvements are accounted for in the cost–benefit method used in Sweden, as well as by other countries and organizations (*19*).

THEORETICAL AND MODELING FRAMEWORK

The authors of the present study chose to estimate the value of freight time indirectly because it is impossible to measure the value of freight time directly from the market. Two methods for estimation of the value of freight time were developed. One method used RP data and the other used SP data. Both methods assume that all owners of goods maximize their utilities.

RP data have the advantage of accurately depicting the freight transportation situation, but they also have the disadvantage of requiring large quantities of data, resulting in the need for long and costly surveys. When the model is set up, the questionnaire may also fail to capture respondent characteristics that are important for the model. In addition, this approach assumes that each person clearly understands the choice posed in the model, whereas in reality, the choice may not be accurately presented to all participants because the decision-making processes for freight operations are complex. The estimate will be biased if the model presents the wrong choice set. In the application of the SP method described here, instead of directly asking the respondents about their willingness to pay, the preferences of a respondent to a questionnaire were analyzed indirectly on the basis of a model and were used to estimate willingness to pay.

To estimate the value of freight time savings, a random utility model based on a choice concept developed and applied to freight

mode choice is introduced. The traditional multinomial logit (MNL) method can be approximated by the maximum likelihood method and is useful for the comparison of results with those from similar previous studies. Although the MNL method is perfectly adequate for many applications, the properties of the MNL method may, in some contexts, become overly restrictive because the MNL method ignores the heterogeneity of economic agents: it assumes that the unobserved error component of all observations are independent and makes the famous independence of irrelevant alternatives (IIA) assumption. The IIA assumption is the most restrictive assumption of the MNL method: it states that a change in the attributes of one alternative changes the probabilities of the other alternatives in proportion. This substitution pattern may not be realistic in all settings. The coefficients of all attributes are also assumed to be the same for all respondents in a choice experiment, whereas in reality there may be substantial variability in how different individuals respond to these attributes (*20*).

The method proposed for use for analysis of the value of rail freight travel time savings is based on the random parameter logit (RPL) model, which allows the data more freedom to directly reveal the form of any inherent taste variation and correct repeated choices bias in the *t*-ratios of parameters estimated by the traditional MNL method (*21*). A shipper's or producer's choice among J alternatives can be analyzed by use of the RPL model. Assume that the shipper's or producer's utility from each alternative can be decomposed into a nonstochastic, linear-in-parameters part that depends on observable variables, a stochastic part that is normally distributed and potentially correlated and heteroscedastic, and another stochastic part that is independently and identically extreme value distributed. Given these assumptions, the utility (U) of individual i from alternative k in choice situation t is denoted

$$U_{nit} = \beta x_{nit} + [\eta_{ni} + \epsilon_{nit}] \tag{1}$$

where

x_{nit} = vector of observed nonstochastic variables, including the socioeconomic characteristics of individual n ($n = 1, \ldots, N$) and the attributes of alternative i in choice situation t;
β_i = I vector of structural parameters;
η_{ni} = error term that is normally distributed over individuals and alternatives; and
ϵ_{nit} = an extreme-value-distributed error term that is independently and identically distributed over individuals, alternatives, and choices by the same person.

The mixed logit model in Equation 1 can alternatively be interpreted as an RPL model in which η_{ni} is the stochastic part of a randomly distributed alternative specific constant. The difference between the two models is entirely based on interpretation. The random parameter specification corresponds to the Limdep programming code used in the estimations for this paper (*22*). The density of η is represented by $f(\gamma/\Omega)$, where Ω represents the fixed parameters. From a given η, the IIA assumption holds, and the conditional choice distribution is the traditional MNL method, since the error terms are IIA with an extreme value distribution.

$$L_{ni}(\eta) = \frac{e^{\beta_i x_{ni} + \eta_{ni}}}{\sum_{j \in J} e^{\beta_j x_{ni} + \eta_{ni}}} \tag{2}$$

where L_{ni} is the conditional choice probability.

Because η is not given, the unconditional choice probability is the traditional logit, which is integrated over all values of η, with the density function of η given as

$$P_{ni} = \int p_{ni}(\eta) f(\eta|\Omega) d\eta \tag{3}$$

where P_{ni} = probability that individual n chooses alternative i.

This choice probability cannot be calculated exactly and is approximated through simulation (*23*).

SURVEY DESIGN AND DESCRIPTIVE ANALYSIS OF DATA

The leading 51 companies that mainly use rail for their freight transportation operations were chosen as the survey sample, as these companies together accounted for 85% of the total rail freight volume of South Korea in 2006. The authors contacted the people responsible for the rail freight movements of the producers, such as freight managers and the executive directors of rail freight sections, as well as individuals from rail freight shippers, including freight forwarders and rail freight operators. The SPs of these individuals were investigated by using a specially designed survey.

The questionnaire was divided into three parts. The first part contained general information about the characteristics of the company (such as the location, annual sales, value of freight, average length of rail freight, and types of commodities). Table 1 presents the descriptive statistics of the sample. About 53% of the respondents had freight operations as their major area of business, with this category including third-party logistics firms. Sixty-two percent of the sample had annual sales of over $50 million. The average length of rail haul was just over 300 km, which is a relatively short distance compared with the distances described in previous studies, such as those from Sweden and Norway. The short distance over which freight is moved by rail is not surprising, considering the size of South Korea. More than 55% of the rail freight volume represented export–import traffic, and most companies used containers as their major mode of rail freight operation. The second part of the survey

TABLE 1 Descriptive Sample Statistics

Variable	Definition	Mean	Standard Deviation
Company	Main business of firm Producer = 1 Shipper = 2	1.4750	0.4995
Sales	Annual sales Below $10M = 1 $10M–$50M = 2 Over $50M = 3	2.4989	0.7081
Commodities	Major commodities Bulk = 1 Container = 2	1.5750	0.4945
Operations	Types of operations Domestic freight = 1 Export–import = 2	1.3250	0.4685
Distance	Average distance Below 200 km = 1 200–400 km = 2 Over 400 km = 3	2.1000	0.3001

NOTE: M = million.

asked the respondents to make a choice from the selected choice sets. The choice sets were based on hypothetical changes in freight operations that an investment in the rail infrastructure could enable. This step determined preferences in terms of changes to the transit time, service reliability, and the frequency of rail service. Various attributes were modeled. The typical attributes used in this type of study were transit time for freight movements, service reliability, service flexibility, service frequency, and the ability to schedule movements at the desired times (*23*).

The second part of the survey was also designed to take into account rail freight service both at present and after the implementation of the proposed changes. Each producer or shipper chose three alternatives for each characteristic. The first option was the present situation, and the second and third options represented the alternatives chosen by the respondent in response to changes to freight transportation by rail. The change in the level of each explanatory variable is shown in Table 2. The study asked 16 questions about the transportation change in a second step, and these questions investigated the producer's or shippers' preferences. The questions were designed to present the respondent with choice conditions, choice alternatives, and explanatory variables and to yield a choice set that reflected the producer's or the shippers' preferences.

The third part of the survey presented questions with variables that influence the alternative choices, excluding any biases. The responses in this step help support the decision on how to weight disregarded or considered variables. An example of an experiment is shown in Table 3.

Most respondents were presented with three alternatives, including current service attributes. The inclusion of the current service attribute reinforces the reality of the proposed alternatives. Only two choices were presented to some respondents. In those cases, an SP alternative was used to limit the effect of respondent fatigue. The SP choices were selected from the full factorial design, which included all possible combinations among the attribute levels. Each of the four attributes had four response levels, resulting in a complete factorial involving 4^4, or 256 combinations. To reduce the magnitude of the task facing the respondents, a common approach was to select the smallest orthogonal main effects plan sampled from the complete factorial design to select the profiles to be used. The smallest orthogonal plan was created by using SPSS software. Under the assumption that all interaction effects are negligible, with the help of SPSS software, the fractional factorial orthogonal scenarios for main effects only were designed. The scenarios yielded 16 alternatives that used all the attributes and levels. However, asking a respondent to choose from among 16 alternatives seemed impossible. To reduce the difficulty of this task for the respondents, two alternatives were randomly selected from the 16 alternatives. Each respondent also faced only one of the three choice sets. These features helped achieve the goal of structuring the experiment to be of as little of a burden as possible without losing the structural properties of the model.

TABLE 2 Service Quality Attributes and Levels

Quality Attribute	Levels
Cost	−10%, current, +10%, +20%
Transit time	−10%, current, +10%, +20%
Reliability	Current, 1-h delay, 2-h delay, 3-h delay
Service frequency	Current, one more trip per day, two more trips per day, three more trips per day

TABLE 3 Typical Choice Example Scenarios for Alternative Attributes

Attribute	Current Services	Alternative 1	Alternative 2
Cost	Same as now	10% more	Same as now
Transit time	Same as now	20% more	10% less
Reliability	Same as now	1 h of delay	2 h of delay
Service frequency	Same as now	Same as now	+2 rail freight trips per day
Choice: (only one)	❑	❑	❑
Choice: (only one)		❑	❑

The MNL and RPL models were used for the various groups, such as producers versus shippers, general cargo versus container, inbound versus outbound cargo, and domestic versus import–export cargo. Most previous studies have relied on the respondents' answers concerning transportation times and rates. However, the government operates the South Korea freight railroad and sets its prices. Therefore, this study determined rail transportation transit times from the freight rail schedule available for each station and rates based on the calculations of the South Korea Logistics Information System, based on the official published rates. The general cargo rate is based on a carload value of \$0.042/ton km and about \$0.876 per 40-ft container kilometer.

This questionnaire was administered from February 1 to March 8, 2006. The respondents agreed to take the questionnaire on February 1 and 2, 2006, and its distribution began on February 3, 2006. A total of 43 questionnaires were collected by March 6, 2006, which corresponds to a response rate of 82%. The high response rate shows that shippers are interested in rail transportation issues. The data based on the collected questionnaires were encoded, and data errors were corrected by phone or fax.

ESTIMATION RESULTS

This section presents the main estimation results obtained with the MNL and RPL models. Only a limited portion of the numerous possible scenarios are intentionally presented. Because the SP data reflect the respondents' preferences under hypothetical conditions, other measures are needed to validate the results. It is important to understand the current state of the transportation system. It is then assumed that the future system incorporates the proposed changes. To minimize bias in the analysis, the choice set disregarded by the greatest number of respondents to the questionnaire was analyzed, and the results are shown below:

Variable	*Frequency (no. of companies)*
Change in costs	10
Change in transit time	8
Change in reliability of departure and arrival times	7
Change in service frequency	15

The choice set selected by the lowest number of respondents (37.5%) was the change in service frequency. It was concluded that changing the frequency of rail service does not influence the estimation of the per hour value of freight time. This is because shippers use rail service an average of only once per day, almost all transportation

TABLE 4 Estimation Results from the MNL and RPL Models for Sample

	MNL-1 (*t*-stats.)	MNL-2 (*t*-stats.)	RPL-1 (*t*-stats.)	RPL-2 (*t*-stats.)
Fixed coefficients				
Transit time	–0.256 (–3.0)	–0.266 (–3.1)	–0.305 (–2.9)	–0.313 (–3.1)
Cost (price)	–0.030 (–8.8)	–0.137 (–4.2)	–0.029 (–8.2)	–0.136 (–4.1)
Reliability	–0.309 (–3.3)	–0.311 (–3.2)	–0.365 (–4.5)	–0.378 (–4.5)
Service frequency	0.000 (0.1)	0.007 (0.2)	0.001 (0.3)	0.014 (0.2)
Random parameter				
Transit time			1.130 (1.7)	1.112 (1.9)
Interaction terms				
Sales*cost		0.004 (1.6)		0.004 (1.5)
Volume*cost		0.003 (2.1)		0.003 (1.9)
Commodity*time		0.008 (0.8)		0.009 (0.9)
Goods*time		0.029 (2.8)		0.029 (2.8)
Number of sample	360	360	360	360
Log likelihood at zero	–395	–395	–395	–395
Log likelihood function	–307	–280	–300	–275
Pseudo-R^2	.22	.24	.26	.28

occurs at night, and the time required for transportation is short because of the geographical setting.

The most statistically significant results of the sample data are presented in Table 4. Table 4 provides the coefficients of the traditional MNL-1 model (standard MNL model), the MNL-2 model (the MNL model with interaction terms), the RPL-1 model (the standard RPL model in which the attributes are normally distributed), and the RPL-2 model (with the interaction terms). In the RPL model, the transportation time variable is defined as random parameters with triangular distributions. Rail freight preferences for transit time may differ according to the respondent's status as either shipper or producer or according to company characteristics, such as annual sales, volume carried, types of commodities, and direction of freight (inbound or outbound).

The values for the basic MNL-1, extended MNL-2, basic RPL-1, and extended RPL-2 models can be estimated by assuming that the random coefficient varies with transit time. Table 4 shows the results of the estimates. The second and third columns present the results of the MNL-1 and MNL-2 models. The results of the chi-square statistic based on the log likelihood ratios allow rejection of the null hypothesis. The basic MNL-1 model is estimated without reference to the economic characteristics of the shippers or the producers, so the estimate depends only on the transit time, cost, reliability, and service frequency variables. All fixed coefficients of the MNL-1 model have the expected sign, and all are statistically significant except for the service frequency variable. Reliability turns out to be the attribute with the most significant effect on the cost of movement because of the time value of freight.

For the extended MNL-2 model, the interaction terms are significantly related to transit time and cost. A log likelihood ratio test leads to rejection of the null hypothesis of the similarity between the MNL-1 and MNL-2 models and demonstrates that addition of new interaction variables changes the model and allows it to explain more heterogeneity. As far as the adjusted R^2 is concerned, the extended MNL-2 model explains the rail freight choice alternatives made by respondents more comprehensively. Therefore, the extended MNL-2 model is statistically superior to the MNL-1 model. As far as the RPL-1 and RPL-2 models are concerned, if the estimated standard deviations of the coefficients are significant, then the parameters of the samples can be said to vary. In this case, the *t*-statistics of the random parameters should be significant, and accordingly, a log likelihood ratio test rejects the RPL models in favor of the MNL-1 and MNL-2 models (*23*).

Columns 4 and 5 of Table 4 present the results for the RPL-1 and RPL-2 models. In the context of this study, the incorporation of heterogeneity by rendering attributes systematically dependent on economic variables such as time, cost, service reliability, and service frequency of rail freight choice does not lead to specifications with more explanatory power than that achieved with the traditional MNL-1 and MNL-2 models. In particular, the RPL-1 and RPL-2 models are not structurally different from the MNL-1 and MNL-2 models when heterogeneity is incorporated. In accordance with these methodological and statistical estimation results, the value of freight travel time saving is presented on the basis of the RPL-2 model.

Table 5 presents the value of rail freight transit time savings for different segments of the sample. When all shipments are considered together, the value is estimated to be about \$2.31/ton h, and the value of rail freight reliability is \$2.78/ton h of delay. The estimated

TABLE 5 Value of Rail Freight Transit Time for Different Sample Segments

Attribute	Transit Time	Service Reliability
Model		
MNL-2	1.95	2.28
RPL-2	2.31	2.78
Producers	1.21	2.27
Shippers	2.35	2.49
Domestic	1.24	2.07
Export–import	2.63	2.84
General cargo	2.23	3.19
Bulk cargo	1.77	2.53
Inbound	1.82	2.98
Outbound	2.03	2.00
Short-distance	2.37	2.77
Long-distance	1.32	2.12
	\$/ton h	\$/ton h of delay

TABLE 6 Number of Preliminary Feasibility Projects Since 1999 (South Korea)

Year	Highway	Railroad	Seaport	Airport	Water–Dam	Other
1999	9	2	1	0	1	4
2000	11	7	5	1	1	5
2001	20	14	1	1	0	5
2002	9	9	0	0	4	7
2003	8	7	3	0	5	7
2004	17	13	0	0	3	15
2005	10	6	2	0	0	9
2006	29	7	4	0	0	14
Total	113	65	16	2	14	66

value of rail freight transit time in South Korea is higher than that determined in previous studies in other countries (*3*) because the maximum rail freight transit time in South Korea is about 6 to 7 h, whereas some other studies report transit times of at least several days. The SP data generally show that the estimated value of freight time per hour varies greatly on the basis of the amount of time savings. This means, for example, that a shipper who is willing to pay \$2/ton to save 1 h of transit time may be willing to pay significantly more or less than \$10/ton to save 5 h of transit time.

The breakdown of the results according to economic agents yields some interesting findings. Namely, shippers (mostly freight forwarders) perceived the value of freight time and reliability to be much higher than producers did, with time savings values of about \$2.35/ton h for shippers and merely \$1.21/ton h for producers. Also, export–import goods are valued relatively greater than domestic shipments, both in terms of the value of rail freight savings and in terms of reliability. Not surprisingly, the perceived values of time and reliability for container traffic are higher than those for bulk cargo. For shipment directions, the values of freight transit time savings and reliability were quite similar between inbound and outbound cargoes. Finally, the value of transit time savings and reliability is inversely related to the length of the haul.

The results of the values of rail freight time and service reliability were applied to the preliminary feasibility studies for the freight railroad investment projects in South Korea, as shown in Table 6. Among 65 railroad projects that required a preliminary feasibility study, 20 projects heavily involved freight movements. Only four of the 20 freight railroad investment projects had benefit–cost ratios greater than 1.0. The remaining 16 projects had benefit–cost ratios ranging from 0.24 to 0.96. When the estimation results were applied to the preliminary feasibility of freight railroad projects in South Korea, it was found that of the 20 rail freight projects, seven projects that had been considered infeasible became feasible. In other words, the inclusion of the value of freight time in the calculation of benefits increased the number of feasible rail freight projects from four to 11, or from 20% to 55%.

CONCLUSIONS

This paper presents an empirical analysis of the discrete choice model for the value of rail freight transit time, reliability, and service frequency in South Korea. The MNL and RPL models were estimated with interaction terms. The MNL models, which assume homogeneity, and the RPL model are both helpful. However, the methodological advantages of the RPL model over the MNL models allow the more accurate and complete determination of the value of freight transit time savings.

When all rail shipments are considered together, a value of \$2.31/ton h is estimated for transit time savings, and the value of rail freight reliability is \$2.78/ton h of delay. The estimated value of rail freight transit time in South Korea is higher than the values determined in previous studies.

The results concerning economic agents are interesting, namely, that the shippers perceived the value of rail freight time and reliability to be higher than producers did. Shippers valued rail freight transit time savings at \$2.35/ton h, whereas for producers this value was merely \$1.21/ton h.

These values were applied to the preliminary feasibility evaluation of 20 freight railroad projects in South Korea. Without consideration of the freight time value, the benefits outweighed the costs for only four (20%) projects. With consideration of the freight time value, the number of projects for which the benefits outweighed the costs increased to 11 (55%).

ACKNOWLEDGMENT

This research was partially supported by the Transport Connectivity and Transfer Facility R&D Program (September 2006) of the South Korea Institute of Construction and Transportation Technology Evaluation and Planning.

REFERENCES

1. *Guidance on Rail Appraisal.* Transport Analysis Guidance (TAG) Unit 3.13.1, Department of Transport, London, Aug. 2007. http://www.webtag.org.uk/webdocuments/3 Expert/13 Rail/pdf/3.13.1.pdf.
2. Massiani, J. The Welfare Effects of Freight Travel Time Savings. Munich Personal RePEc Archive (MPRA) Paper 8754. Munich, Germany, May 2008.
3. Zamparini, L., and A. Reggiani. Freight Transport and the Value of Travel Time Savings: A Meta-Analysis of Empirical Studies. *Transport Reviews,* Vol. 27, No. 5, 2007, pp. 621–636.
4. Fowkes, T. The Design and Interpretation of Freight Stated Preference Experiments Seeking to Elicit Behavioral Valuations of Journey Attributes. *Transportation Research, Part B,* Vol. 41, 2007, pp. 966–980.
5. Transek, A. B. Godskunders Transportmedelsval. Report from the project Samord nat trafikanalyssystem i Malardalen (SAMM), Göteborg, Sweden, May 1992.
6. Fridstrom, L., and A. Madslien. *A Stated Preference Analysis of Wholesalers? Freight Choice.* TOII Report 1995/09. Transport Economics Institute, Oslo, Norway.
7. Wynter, L. M. Stated Preference Survey for Calculating Values of Time of Road Freight Transport in France. In *Transportation Research Record 1477,* TRB, National Research Council, Washington, D.C., 1995, pp. 1–6.
8. Wynter, L. M. The Value of Time of Freight Transport in France: Estimation of Continuously Distributed Values from a Stated Preference Survey. *International Journal of Transport Economics,* Vol. 22, 1995, pp. 151–165.
9. Jeong, I. Y. The Effect of Road Investment on Logistics Cost in Manufacturing Industry. *Journal of Korean Society of Transportation,* Vol. 17, No. 1, 1993, pp. 12–20.
10. Kurri, J., A. Sirkia, and M. Juha. Value of Time in Freight Transport in Finland. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1725,* TRB, National Research Council, Washington, D.C., 2000, pp. 26–30.

11. Kawamura, K. Perceived Value of Time for Truck Operators. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1725,* TRB, National Research Council, Washington, D.C., 2000, pp. 31–36.
12. Wigan, M., N. Rockcliffe, T. Thoresen, and D. Tsolakis. Valuing Long-Haul and Metropolitan Freight Travel Time and Reliability. *Journal of Transportation and Statistics,* Vol. 3, No. 3, 2000, pp. 83–89.
13. de Jong, G. Specification of Logistics Model for Norway and Sweden. Presented at 85th Annual Meeting of the Transportation Research Board, Washington, D.C., 2006.
14. Hodkins, K. E., and D. N. M. Starkie. Values of Time in Long-Distance Freight Transport. *Logistics and Transportation Review,* Vol. 14, No. 2, 1978, pp. 117–126.
15. Blauwens, G., and E. van de Voorde. The Valuation of Time Savings in Commodity Transport. *International Journal of Transport Economics,* Vol. 15, 1988, pp. 77–87.
16. Viera, L. F. M. *The Value of Service in Freight Transportation.* PhD dissertation. Massachusetts Institute of Technology, Cambridge, 1992.
17. Swait, J., J. Louviere, and M. Williams. A Sequential Approach to Exploiting the Combined Strengths of SP and RP Data: Application to Freight Shipper Choice. *Transportation,* Vol. 21, 1994, pp. 135–152.
18. Kang, W. The Value of Rail Freight Time in Korea. Presented at 7th Meeting of the Eastern Asia Society of Transportation, 2007.
19. Bruzelius, N. *The Valuation of Logistics Improvements in CBA of Transport Investments: A Survey.* Report to the SAMGODS Group. Swedish Institute for Transport and Communications Analysis, Östersund, 2001.
20. Basu, D., and B. Maitra. Valuing Attributes of Enhanced Traffic Information: An Experience in Kolkata. *Transport,* Vol. XXII, No. 3, 2007, pp. 164–173.
21. Brownstone, D., and K. Train. Forecasting New Product Penetration with Flexible Substitution Patterns. *Journal of Econometrics,* Vol. 99, pp. 109–129, 1999.
22. Greene, W. H. *Limdep Version 7.0 Users Manual.* Econometric Software, Inc., New York, 1998.
23. McFadden, D., and K. Train. Mixed MNL Models of Discrete Response. *Journal of Applied Econometrics,* Vol. 15, 2000, pp. 447–470.

The Local and Regional Rail Freight Transport Committee peer-reviewed this paper.

Effect of Train Speed on Risk Analysis of Transporting Hazardous Materials by Rail

Athaphon Kawprasert and Christopher P. L. Barkan

This paper considers the effect of train speed on railroad hazardous materials transportation risk. A statistical method was developed to estimate the speed-dependent conditional probability of release (CPR) of hazardous material from tank cars involved in accidents. The objective was to assess how accounting for speed affects the results of risk analysis. A case study of a representative hazardous materials route was conducted. In that case study, risk estimates were developed by using conventional, average-speed CPR, and the estimates were compared with those obtained by analyzing the same route using speed-dependent CPR. The differences in calculated risk indicate that the use of the speed-dependent CPR may be an important refinement for the accurate calculation of the risk of a route. The effects of track-class upgrades on risk were also considered. Track with higher Federal Railroad Administration classes has lower accident rates, but the higher permissible speeds increase the CPR if a tank car is derailed in an accident. Consequently, evaluation of the effect of track upgrades must account for both factors to understand the net impact on risk. Two scenarios were considered in the track-class upgrade analysis: upgrades without an increase in speed and upgrades with an increase in speed. The increased tank car hazardous material release rate at the higher speeds permissible on higher track classes was more than offset by the reduction in the accident rate for all the track-class upgrade scenarios considered. For the particular route analyzed in the case study, use of the speed-dependent CPR resulted in a slight increase in the overall risk estimate, and upgrading of Class 3 track provided the greatest reduction in risk. Such results will be route specific; however, the use of a speed-dependent CPR enables the more accurate analysis of local risk and provides a better evaluation of risk reduction options that involve changes in operating and track characteristics.

Train operating speed is one of the factors affecting the likelihood of a release in railroad accidents involving tank cars transporting hazardous materials. With a higher speed of derailment, more cars are likely to derail, and for those derailed cars there is a higher probability that one or more will suffer a release (*1*). Nevertheless, previous studies have generally simplified this aspect of railroad hazardous materials transportation risk analysis by using a single value for the conditional probability of release (CPR) that does not account for variations in train speeds (*2–5*). Use of a single value of CPR, independent of the train speed, simplifies the risk analysis but also implies an average speed of derailment (*6*). Because this average-speed CPR may be higher than the actual operating speed in some situations and lower in others, use of the average-speed CPR will have the effect of overestimating the probability of release on lower-speed sections and underestimating the probability of release on higher-speed sections of the route.

The relationship between speed and CPR has been considered previously (*7–9*). The most recent study (*7*) provides estimates for both the average-speed CPRs and the speed-dependent CPRs of tank cars, but only the average-speed CPR accounts for specific safety design features. Earlier work was not specific to tank cars (*8*) or did not enable the application of speed-dependent effects to specific tank car designs (*9*). In the study described here, the authors explicitly considered the effects of train speed on hazardous materials transportation risk analysis and developed a technique that enables estimation of the speed-dependent CPR for specific tank car designs using published data on the safety performance of tank cars in accidents (*7*) that permits the application of adjustment factors derived from a group of car types to a specific car type.

The use of tank car design enhancements as options to reduce risk has become more prevalent over the past decade, and the model introduced here provides flexibility that enables the effects of these enhancements to be considered by using the most up-to-date published data available. Furthermore, the effect of tank car safety design can be integrated with other options.

In this paper, the authors specifically consider the interaction of infrastructure quality, train speed, and tank car performance to understand the effect on risk estimates using the speed-dependent CPR. A case study that uses a representative hazardous material transportation route is presented. The results of a risk analysis that uses the average-speed CPR are compared with those of a risk analysis that uses the speed-dependent CPR.

The second part of the paper highlights the importance of using CPR adjusted for speed by further considering its utility as part of an assessment of the effect of an upgrade in the track class on risk. Upgrading the track class has been shown to be correlated with the reduction of certain types of accidents and, consequently, risk (*8*). However, if the upgrade is also intended to allow increased operating speeds, it may increase the probability of release if an accident does occur (*1, 7*). The CPR and the accident rate are the two principal elements in hazardous materials transportation risk analysis. The consideration of each one as a function of speed enables the more accurate estimation of route-specific risk. This also facilitates proper consideration of the benefit of infrastructure improvement. Two scenarios were analyzed: a track upgrade without a speed increase and a track upgrade with a speed increase. The degree of risk reduction varies with these options, and consideration is given to which option offers the greatest safety benefit for the problem analyzed.

A. Kawprasert, 3214 Newmark Civil Engineering Laboratory, and C. P. L. Barkan, 1203 Newmark Civil Engineering Laboratory, MC-250, Department of Civil and Environmental Engineering, University of Illinois at Urbana–Champaign, 205 North Mathews Avenue, Urbana, IL 61801. Corresponding author: A. Kawprasert, akawpra2@illinois.edu.

Transportation Research Record: Journal of the Transportation Research Board, No. 2159, Transportation Research Board of the National Academies, Washington, D.C., 2010, pp. 59–68.
DOI: 10.3141/2159-08

REVIEW OF RISK ANALYSIS METHODS

Hazardous materials transportation risk can be quantitatively expressed as the frequency of a release incident multiplied by the consequence of that release. In this study, frequency is the product of the annual rate of tank car involvement in an FRA–reportable derailment on main-line track, the number of shipments, the total mileage from the origin to the destination, and the CPR given that a tank car derails. The consequence of a release incident is the impact of the released material and is affected by the characteristics of the released product, its quantity, and the rate of spillage; atmospheric conditions; and the population density along the route analyzed.

The consequence can be expressed by using several metrics. In the present study, the authors used the number of people who might be affected because of the release of a hazardous material from a tank car, in accordance with the product-specific recommendations presented in the U.S. Department of Transportation (USDOT) *Emergency Response Guidebook* (ERG) (*10*) and elsewhere (*11*). The consequence can be estimated by multiplying the area affected by using the evacuation distances recommended in ERG by the population density within the affected area.

The risk metric calculated by using the method described above is the annual expected number of people who might be evacuated or sheltered in place because of the release of hazardous materials. Specifically, risk is calculated from the product of the accident rate, the traffic volume (measured in car miles), the tank car CPR, the probability distribution of release sizes, the affected area corresponding to the release size, and the population density in the affected area summed over all segments on the route.

The risk model in Equation 1 incorporates these segment-specific parameters and gives the estimate of risk associated with shipments of hazardous materials on the route.

$$S = \sum_{j=1}^{\Pi} \sum_{j=1}^{k} Z_i V_i L_i R_i P_j A_j D_i \tag{1}$$

where

S = annual risk (number of people affected per year);
Z = rate of tank car involvement in an FRA-reportable derailment (number of cars derailed per car mile);
V_i = annual shipments (number of carloads per year);
L_i = length of track segment (miles);
R_i = CPR given that a tank car derails;
P_j = CPR of specific scenario j given that there is a hazardous material release from a derailed tank car;
A_j = affected area corresponding to a specific release scenario, according to the USDOT ERG recommendation (square miles);
D_i = average population density along track segment (number of people per square mile);
i = track-specific segment;
n = total number of segments along the route considered;
j = specific release scenario (e.g., spill size, fire involvement, and time of day); and
k = total number of scenarios considered.

CASE STUDY

A typical, representative route that is part of the distribution network of a particular hazardous material on the North American rail network was considered. The distance from the origin to the destination of the route is 1,400 mi, the route comprises 598 segments, and the average segment length is 2.34 mi. These track segments correspond to links in the FRA rail transportation network obtained from the Bureau of Transportation Statistics (*12*). The authors assumed 100 annual carloads over the entire route using a noninsulated, USDOT-111A100W1 tank car with 7/16-in. tank thickness and no special safety design features beyond the USDOT minimum requirements.

For some materials, consideration of multiple release scenarios may be appropriate. However, on the basis of expert opinion of the manufacturer of the product and the objectives of the present study, use of a single release scenario was satisfactory. Thus, the major factors affecting risk in this analysis were the product characteristics, the track-class-specific accident rate, the shipment volume, the safety design of the tank car, the operating speed, and the population exposure along the route.

The case study is used to analyze the effects of train speed on the risk estimate for the route and the contributions of various other factors to the risk, including population density and FRA track class. In addition, the case study is used to illustrate the merit of using the speed-dependent CPR in an assessment of the effect of a track infrastructure upgrade on risk reduction. These will be discussed in more detail in the subsequent sections.

ESTIMATION OF PARAMETERS AFFECTING RISK

FRA track-class-specific accident rates were used in this study (*13*). The track speed reflects the FRA track class, which has been shown to be correlated with railroad accident rates (*8*). The authors used railroad timetable speeds to infer the FRA track class and other local operating restrictions for all segments along the route. The next step was to estimate the CPR given that a tank car is derailed in an accident. The relationships developed by Treichel et al. were used to determine the CPR of the particular design of tank car considered in this study (*7*). Because CPR is also affected by train accident speed (*1, 7–9*), the authors adjusted it according to the timetable speed for each track segment. The authors calculated the speed-dependent CPR using the procedure described in the following section.

To estimate the consequences of a release, the authors used geographic information system software, ArcGIS Desktop, Version 9.2, to create the shipment route using the USDOT national rail network (*12*). An overlay analysis of the population distribution along the rail network was conducted by using census tract data from the Environmental Systems Research Institute, Inc., (ESRI) Data and Maps electronic database (*14*), and a buffer representing the exposure area was created. This was the area within the radius from the track center equal to the USDOT ERG maximum evacuation distance for the material considered. The average population density of the affected area corresponding to each track segment was then determined.

The procedure used to estimate and measure risk is similar to that described in previous studies (*2, 6*), except that this study focuses on refining the method to understand the effect of the speed-dependent CPR.

SPEED-DEPENDENT CONDITIONAL PROBABILITY OF RELEASE

Treichel and colleagues described the lading loss probability on the basis of the four major cause-specific loss events as (*7*)

$$P = P\left(E_h \cup E_s \cup E_i \cup E_b\right) \tag{2}$$

where

P = probability that a tank car released its contents given that it derailed in an accident,
E_h = event in which the contents were lost because of head damage,
E_s = event in which the contents were lost because of shell damage,
E_t = event in which the contents were lost because of top fitting damage, and
E_b = event in which the contents were lost because of bottom fitting damage.

The cause-specific loss events have low correlations (*7*), so the authors assumed that E_h, E_s, E_t, and E_b are independent. Therefore,

$$\begin{aligned} P &= 1 - P^c \\ &= 1 - P\left(E_h^c \cap E_s^c \cap E_t^c \cap E_b^c\right) \\ &= 1 - P\left(E_h^c\right)P\left(E_s^c\right)P\left(E_t^c\right)P\left(E_b^c\right) \\ &= 1 - \left[1 - P\left(E_h\right)\right]\left[1 - P\left(E_s\right)\right]\left[1 - P\left(E_t\right)\right]\left[1 - P\left(E_b\right)\right] \end{aligned} \quad (3)$$

where c is the complement of an event.

Let R equal the CPR given that a tank car derails in an accident, unadjusted for speed, and that R_h, R_s, R_t, and R_b are the CPRs, unadjusted, attributed to head, shell, top fitting, and bottom fitting, respectively, so that

$$R = 1 - \left(1 - R_h\right)\left(1 - R_s\right)\left(1 - R_t\right)\left(1 - R_b\right) \quad (4)$$

where

R = tank car CPR, unadjusted;
R_h = CPR from head, unadjusted;
R_s = CPR from shell, unadjusted;
R_t = CPR from top fittings, unadjusted; and
R_b = CPR from bottom fittings, unadjusted.

The speed-dependent CPR was calculated by multiplying the unadjusted CPR by the speed adjustment factors. This adjustment to the CPR associated with each specific source of release can be made by use of Equation 5:

$$R' = 1 - \left(1 - R_h J_h\right)\left(1 - R_s J_s\right)\left(1 - R_t J_t\right)\left(1 - R_b J_b\right) \quad (5)$$

where

R' = tank car CPR, adjusted for speed,
J_h = speed adjustment factor for CPR from head,
J_s = speed adjustment factor for CPR from shell,
J_t = speed adjustment factor for CPR from top fittings, and
J_b = speed adjustment factor for CPR from bottom fittings.

To determine the speed adjustment factors in Equation 5, the authors first developed the relationships between train speed and CPR for each release source by using the proportion of tank cars losing lading from each source. The authors used statistical software (SAS, Version 9.1) to fit a simple linear regression equation with zero intercept to the published data on tank car safety performance (*7*) (Figure 1). A simple linear regression was used without weighting by the number of observations at each speed, consistent with a similar analysis by Treichel et al. (*7*). The fitted functions are shown below, and the corresponding test statistics are summarized in Table 1.

$$Y_h = 0.00786X \quad (6)$$

$$Y_s = 0.00674X \quad (7)$$

$$Y_t = 0.00460X \quad (8)$$

$$Y_b = 0.00150X \quad (9)$$

where

Y_h = proportion of nonpressure cars releasing from heads, corresponding to train speed X;
Y_s = proportion of nonpressure cars releasing from shells, corresponding to train speed X;
Y_t = proportion of nonpressure cars releasing from top fittings, corresponding to train speed X;
Y_b = proportion of nonpressure cars releasing from bottom fittings, corresponding to train speed X; and
X = train speed (mph).

In the next step, the authors used the data from Treichel et al. (*7*) to calculate the weighted average train speeds for releases from the tank head, shell, top fittings, and bottom fittings, which were 38.5, 41.2, 28.7, and 35.5 mph, respectively. These average speeds were then substituted into Equations 6 through 9 to determine the proportion of cars releasing from each source at the average speed (denoted by the subscript *a*), yielding the following values for the noninsulated USDOT-111A100W1 tank car considered here: $Y_{ha} = 0.30260$, $Y_{sa} = 0.27753$, $Y_{ta} = 0.13178$, and $Y_{ba} = 0.05330$, respectively.

The speed adjustment factors were then determined by dividing the proportion of nonpressure cars releasing at a particular speed by the proportion of cars releasing corresponding to weighted average speed, for example, $J_h = Y_h / Y_{ha}$. The speed adjustment factors applicable to nonpressure cars were

$$J_h = 0.02597X \quad (10)$$

$$J_s = 0.02429X \quad (11)$$

$$J_t = 0.03491X \quad (12)$$

$$J_b = 0.02814X \quad (13)$$

The cause-specific CPRs for the particular tank car considered were estimated by using the data from Treichel et al. (*7*) as $R_h = 0.0799$, $R_s = 0.1092$, $R_t = 0.1577$, and $R_b = 0.0625$. By using Equation 4, the average-speed (unadjusted) CPR for this tank car is 0.3527. Substituting the values of cause-specific CPRs into Equation 5, the speed-dependent CPR was estimated as follows:

$$R' = 1 - \left[\left(1 - 0.0799J_h\right)\left(1 - 0.1092J_s\right)\left(1 - 0.1577J_t\right)\left(1 - 0.0625J_b\right)\right] \quad (14)$$

Figure 2 shows the speed-dependent CPRs compared with the average-speed CPR of the noninsulated USDOT-111A100W1 tank car considered. This method can be adapted for any other type of tank car for which suitable data are available (*7*).

EFFECTS OF TRAIN SPEED ON RISK

The authors evaluated the effects of train speed by comparing the risk estimates calculated by using the speed-dependent CPR with those from the baseline case calculated by using the average-speed CPR. Use of the speed-dependent CPR yielded an annual risk of 1.428 people affected per year, whereas the annual risk is 1.291 people affected per year if the average-speed CPR is used, an 11% difference. Further detail regarding the effect of speed-dependent CPR

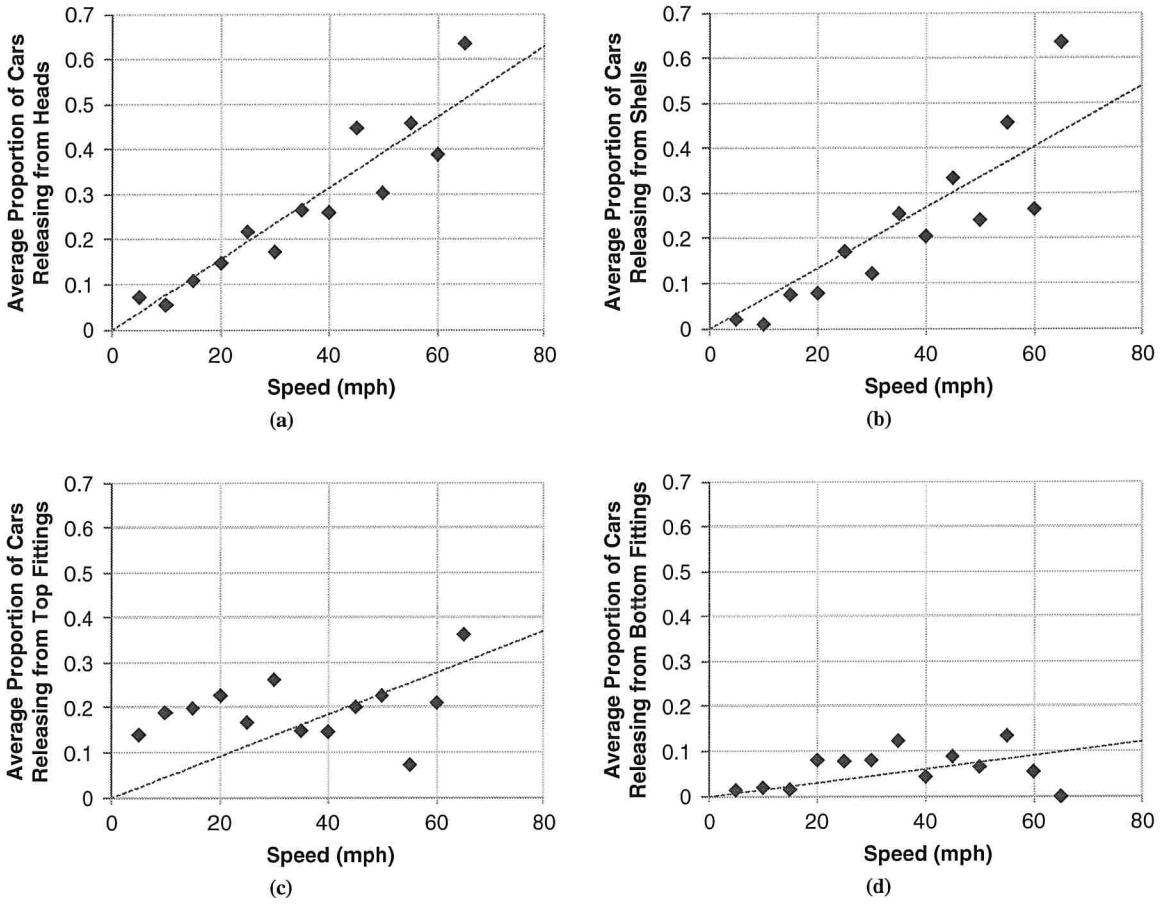


FIGURE 1 Proportion of nonpressure tank cars releasing versus speed for releases from (*a*) heads, (*b*) shells, (*c*) top fittings, and (*d*) bottom fittings.

can be seen by comparing the risk profiles calculated by using the speed-dependent and the average-speed CPRs (Figure 3*a*).

These differences in estimated risk obtained when speed-dependent CPR is used are specific to the characteristics of the particular route analyzed in this case study. In general, the effect on risk estimates will depend on the distribution of speeds along a route. Specifically, routes with a larger percentage of higher-than-average-speed trackage will tend to have increased risk estimates when the speed-dependent CPR is used, and those with a lower percentage will tend to have reduced risk estimates.

In previous work, it was found that a large percentage of the risk along a route was attributable to a small percentage of its length (*6*). In the case study described here, the 100 segments with the highest risk per mile (Figure 3*b*) accounted for 18% of the route length but 92% of the risk. Of these segments, all but 22 had higher estimated risks when the speed-dependent CPR was used than when the average-speed CPR was used. Interestingly, these 22 segments were among the segments with the very highest risk in the entire analysis, accounting for 2% of the route length but 23% of the total risk. In contrast to the overall risk analysis results for the route, use of the speed-dependent CPR for these segments resulted in risk estimates lower than those when average-speed CPR was used.

It is not surprising that use of the speed-dependent CPR resulted in higher overall risk estimates in the case study considered because the segments with higher-than-average speeds comprised a majority of the overall route length, that is, 1% for speeds of 11 to 25 mph (Track Class 2), 16% for speeds of 26 to 40 mph (Track Class 3), 39% for speeds of 41 to 60 mph (Track Class 4), and 44% for speeds 61 to 70 mph (Track Class 5). FRA regulations permit the operation of freight trains up to 80 mph on Class 5 track (*15*); however, 70 mph is a more typical maximum speed and was the case for the route studied. The maximum speed for Key trains, which are subject to a special set of operating practices defined by the Association of American Railroads (*16*), is 50 mph, but the particular hazardous material considered in this study does not affect Key train status. For the purposes of this study, the maximum permissible speeds from railroad timetables were used to infer the FRA track class and were assumed to represent the operating speeds on each track segment.

TABLE 1 Parameter Estimates of Linear-Speed CPR Models

Model	Estimate of Regression Coefficient, β	Standard Error	*t*-Value	Pr > |*t*|	R^2
$Y_h = \beta_h X$	0.00786	0.00044925	17.49	<.0001	.9622
$Y_s = \beta_s X$	0.00674	0.00061609	10.94	<.0001	.9089
$Y_t = \beta_t X$	0.00460	0.00072981	6.30	<.0001	.7677
$Y_b = \beta_b X$	0.00150	0.00031975	4.69	.0005	.6468

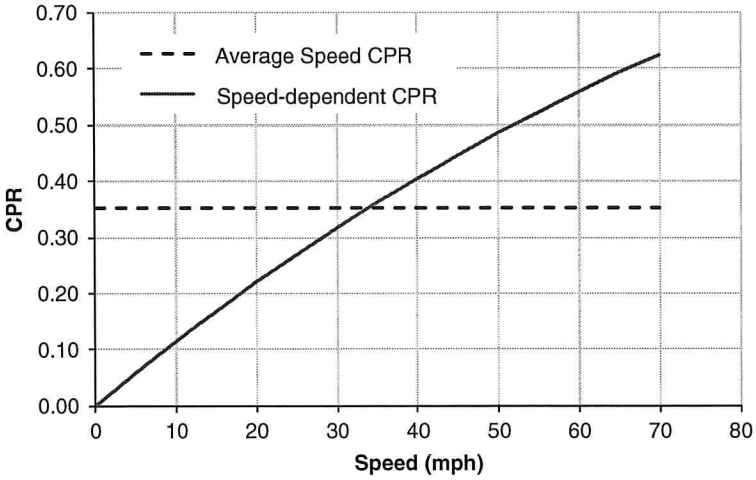
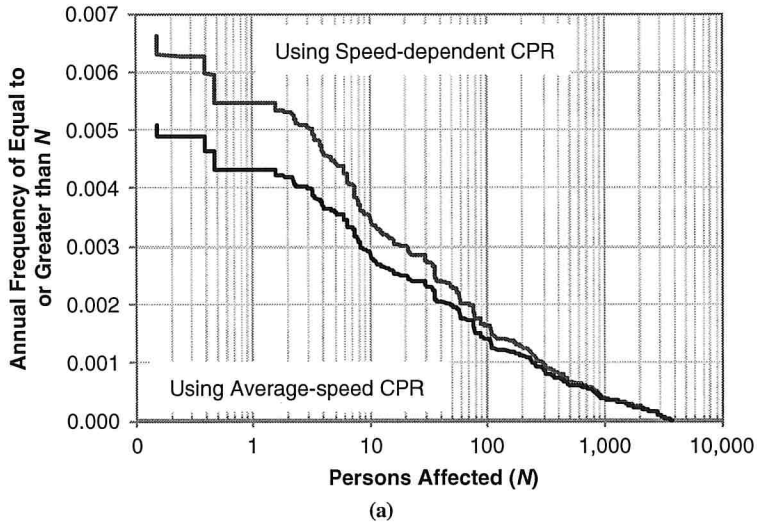
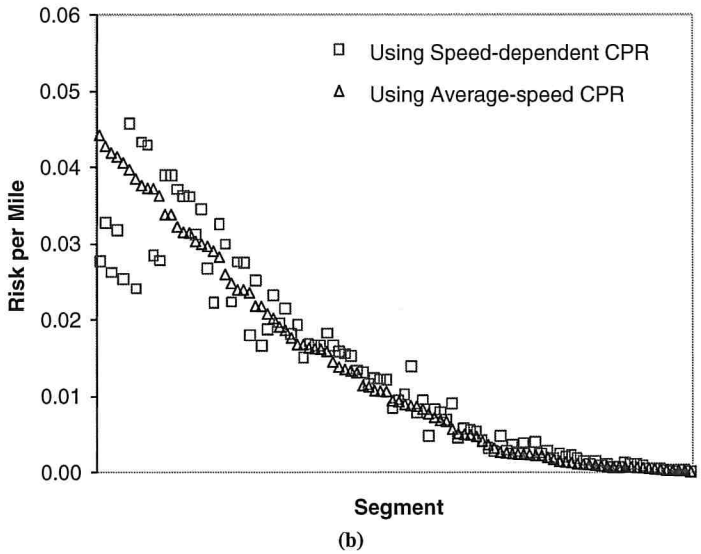


FIGURE 2 Estimated CPR on basis of speed for tank car considered (USDOT-111A100W1).



(a)



(b)

FIGURE 3 Comparison of route-specific results when speed-dependent CPR is used: (*a*) risk profiles (*F-N* curves) and (*b*) top 100 segments with the highest risk per mile.

A potential drawback of using timetable speeds to infer track class is that some segments may actually have higher-quality track and correspondingly lower accident rates than would be assumed on the basis of inferences made from the timetable speeds. In some circumstances, railroads may be maintaining trackage to a higher standard than that required by the track safety regulations for the particular speed shown in the timetable. If information on actual maintenance standards and operating speed is available, the information can be formally incorporated into the risk model, allowing more accurate estimates of accident rates for these segments.

Even in the absence of such information, the use of the speed-dependent CPR instead of the average-speed CPR provides more accurate risk estimates because it better reflects the particular operating characteristics of a route and their effects on risk. Such refinement is particularly important for comparison of the risk associated with different route alternatives, for example, a shorter route passing through urban areas with lower train speeds versus a longer route that goes through less populated areas where trains may operate at higher speeds. Such considerations are germane to railroads and the USDOT, when the results of route risk assessments, as required under HM-232E (*17*), are considered.

RISK CONTRIBUTION BY POPULATION DENSITY AND TRACK CLASSES

The authors analyzed the contributions to risk of different population density groups, using both average-speed and speed-dependent CPRs, to understand their effects on risk estimates with respect to the different population densities for the representative route. Track segments that are located in highly populated areas contribute a large proportion of the route risk. For example, those segments located in areas where the population density exceeds 6,500 people per square mile represent the smallest proportion of the total route length but account for the largest percentage of the risk when the speed-dependent CPR is used (Figures 4*a* and 4*b*).

For the route considered, the use of the speed-dependent CPR yields higher risk estimates (Figure 4*a*, solid line) than the use of the average-speed CPR (Figure 4*a*, dashed line) for all population density groups. The difference between risk estimates for the two CPRs is the largest for the highest-density group and the smallest for the lowest density group. However, when the contribution is expressed as a percentage of the total route risk instead of an absolute estimate; use of the speed-dependent CPR results in a lower contribution for

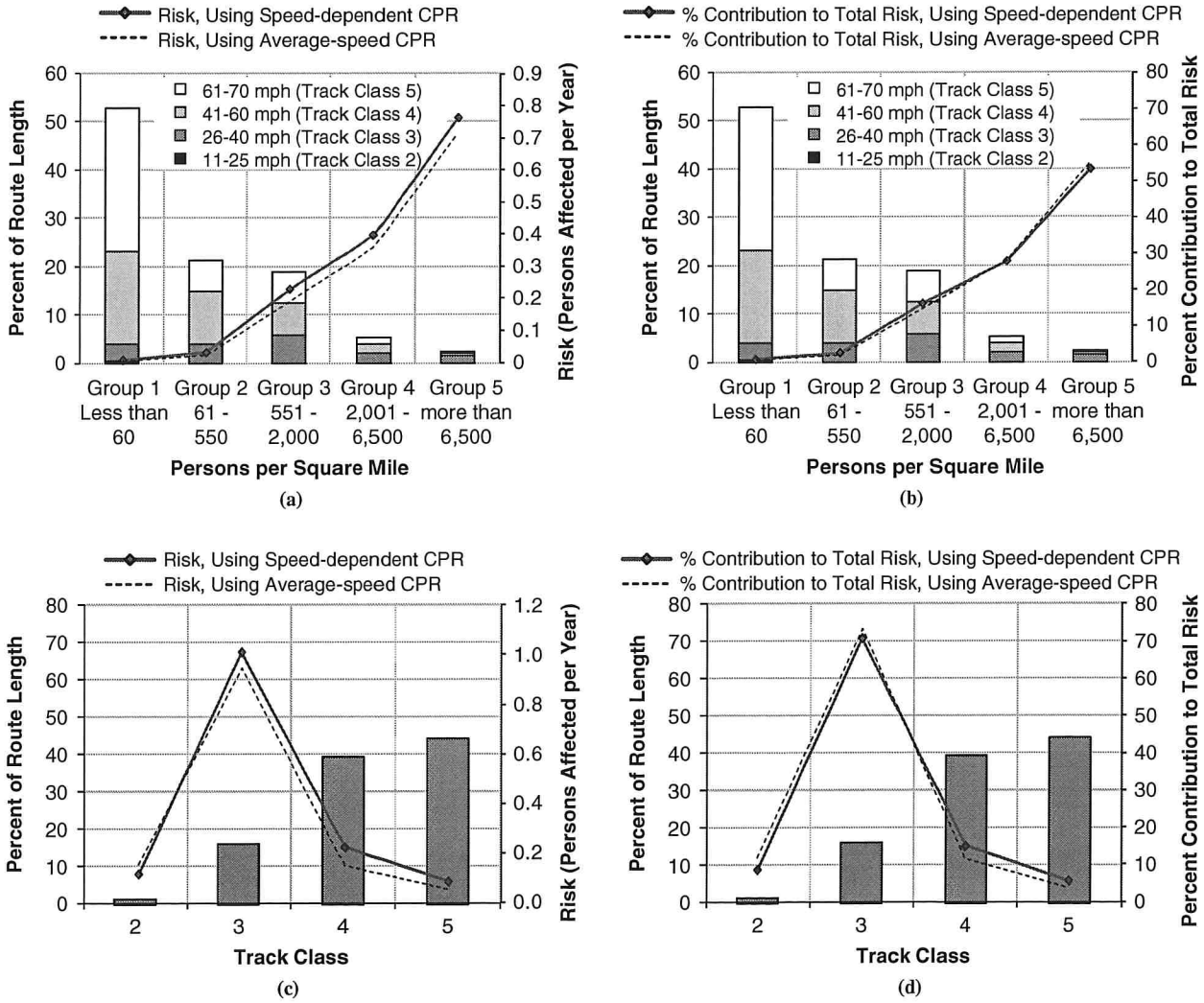

Risk, Using Speed-dependent CPR
Risk, Using Average-speed CPR
61-70 mph (Track Class 5)
41-60 mph (Track Class 4)
26-40 mph (Track Class 3)
11-25 mph (Track Class 2)
Percent of Route Length
Risk (Persons Affected per Year)
Group 1 Less than 60
Group 2 61 - 550
Group 3 551 - 2,000
Group 4 2,001 - 6,500
Group 5 more than 6,500
Persons per Square Mile
(a)
% Contribution to Total Risk, Using Speed-dependent CPR
% Contribution to Total Risk, Using Average-speed CPR
Percent Contribution to Total Risk
(b)
Track Class
(c)
(d)

FIGURE 4 Percentage of route length and contribution to risk on route studied: (*a*) population density versus risk, (*b*) population density versus percentage of total risk, (*c*) track class versus risk, and (*d*) track class versus percentage of total risk.

the highest population density group relative to that obtained by use of the average-speed CPR (Figure 4*b*).

Some urban area segments do not experience the reduction in estimated risk when the speed-dependent CPR is used because in the case study these segments had speeds higher than average. On the other hand, some segments in nonurban areas had lower risk when the speed-dependent CPR was used because of speed restrictions on these segments. Figures 4*a* and 4*b* also show that the majority of the representative route is Class 4 and Class 5 track. Most of these segments are located in areas with moderate to low population densities. Class 3 segments, however, are more evenly distributed over different population classes, from the lowest to the highest density.

The segments with higher track classes do not contribute much risk, despite their greater percentage of route length. By contrast, the segments with lower track classes contribute a larger percentage of the total route risk because of a combination of the accident rate and the population densities associated with them, despite the lower CPRs.

Use of the speed-dependent CPR yields a lower estimate of risk on Class 2 track but higher estimates for Track Classes 3, 4, and 5 (Figure 4*c*). When the percent contribution to total risk is considered (Figure 4*d*), use of the speed-dependent CPR results in a lower percentage of risk on Track Classes 2 and 3 and larger percentages on Track Classes 4 and 5 compared with the values obtained when the average-speed CPR is used. This analysis, based on the representative route, illustrates that use of the speed-dependent CPR takes into account the effect of differential operating speeds on risk for each categorical variable considered: population density, and track class.

In the context of this discussion, it is useful to consider the ratio of the percentage of total risk to the percentage of route length. For the route analyzed, the risk/route length ratios, based on the speed-dependent CPR, were 8.6, 4.4, 0.4, and 0.1 for Track Classes 2, 3, 4, and 5, respectively. Although Class 2 track has the highest risk contribution per track mile, the overall contribution from Class 2 track is not as high as that from Class 3 track because of the small percentage of Class 2 track on the route. Because Class 3 track contributes the largest percentage of risk on this route, measures to reduce this are likely to have the greatest overall effect. In the next section, the authors consider the effects of infrastructure upgrades on risk and illustrate how the use of speed-dependent CPRs enables the better assessment of the safety benefits of track infrastructure upgrades.

EFFECTS OF INFRASTRUCTURE UPGRADES

In the case study, the authors considered the effect of upgrading all Class 2, Class 3, and Class 4 segments to the next higher class, that is, to Classes 3, 4, and 5, respectively, as three distinct upgrade options. The authors first considered the effect of track-class upgrade strictly as a means of reducing the accident rate and, consequently, the risk and therefore assumed that the operating speeds were held constant after the track upgrades. Upgrading Class 2, 3, and 4 tracks to the next higher class yielded reductions in risk of 5%, 53%, and 7%, respectively. Upgrading Class 3 segments offered the greatest reduction in risk, primarily because of the relatively large reduction in the accident rate combined with the fact that Class 3 tracks make up a fairly large percentage of the route analyzed.

Track upgrade principally affects track-caused accidents; however, not all track-related accidents are affected in the same way by track-class upgrades. Furthermore, there are a wide variety of potential causes of accidents (*18*). Non-track-related causes may have no relationship or only an indirect one to track class. Some causes of accidents, notably, certain ones attributable to equipment failure, such as broken wheels or axles, may actually increase with track-class upgrades if the operating speed also increases. Hence, increasing the track class affects the likelihood of only some types of accidents, and the functional relationship between each of these accident causes and the track class varies. When one also considers that different accident causes have different relationships to the likelihood that a derailed car carrying hazardous materials would suffer a release (*1*), it further complicates calculation of the effect of track-class upgrades on risk reduction. Further research on both the statistical and the causal relationships between the track class and accident frequency and severity is needed to improve quantitative assessments of the effects of changes in infrastructure quality on safety and risk. Until such research is completed and comprehensive data on critical infrastructure parameters are generally available, track class remains the best proxy statistic for estimating track quality, the derailment rate, and ultimately, the risk.

COMBINED EFFECTS OF INFRASTRUCTURE AND TRAIN SPEED

This section examines the relationships between track class, accident rate, and speed-dependent CPRs. It was assumed that train speeds increase in accordance with track-class upgrades. That is, trains are assumed to operate at the maximum normal operating speed corresponding to the upgraded track classes. Therefore, for each segment, the accident rate will be reduced because of the track-class upgrade, but the tank car CPR will increase because of the higher speed (Figure 5). The overall release rate, which is the product of the accident rate and CPR (in this example, for a noninsulated USDOT-111A100W1 tank car), is dominated by the accident rate and thus also declines. The difference in the magnitude of the release rate between consecutive track classes is the smallest for Track Classes 4 and 5.

The authors examined each individual segment to understand the change in risk as a result of track infrastructure upgrade and an increase in speed. Figure 6 shows the distribution of the segment-specific risk per mile for three cases for each track class: (*a*) the baseline case (no upgrade), (*b*) upgrade to a higher class without an increase in speed, and (*c*) upgrade to a higher class with an increase in speed. Speed-dependent CPRs were used in all scenarios. For clarity, only the top 10 segments with the highest risk per mile, ordered by the baseline risk, are shown. These charts indicate that upgrading Class 3 segments to Class 4 yields the largest reduction in risk, whereas upgrading Class 4 track to Class 5 offers little reduction in risk for the representative route. Furthermore, the difference between the risk per mile for a segment when speed is held constant compared with the risk when the speed is increased is the highest for the case in which Class 2 segments are upgraded to Class 3. The magnitude of the difference varies because of the different initial operating speeds on each segment (Figure 6*a*). Overall, use of the speed-dependent CPR enables individual consideration of both of the factors affecting risk, in this case, the reduction in the accident rate due to the upgraded track class, and the increase in CPR due to the higher operating speed.

DISCUSSION OF RESULTS

When risk estimates for each of the scenarios analyzed in the case study are summarized, several results can be discerned (Table 2).

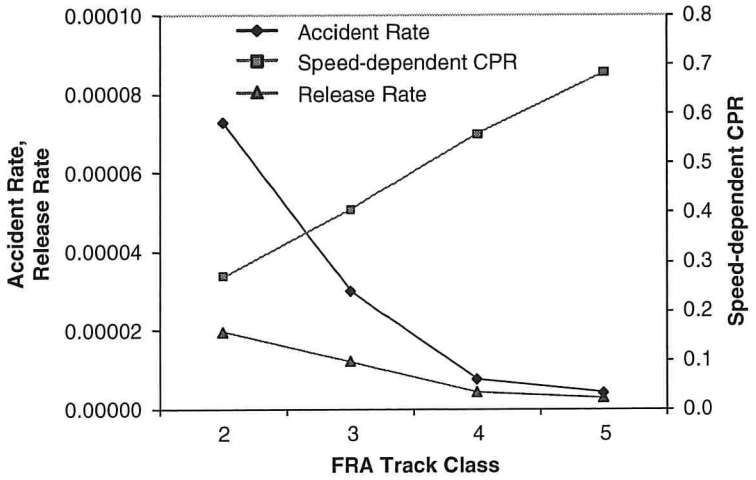


FIGURE 5 Relationship between track classes, accident rates, speed-dependent CPRs, and release rates for tank car considered (USDOT-111A100W1).

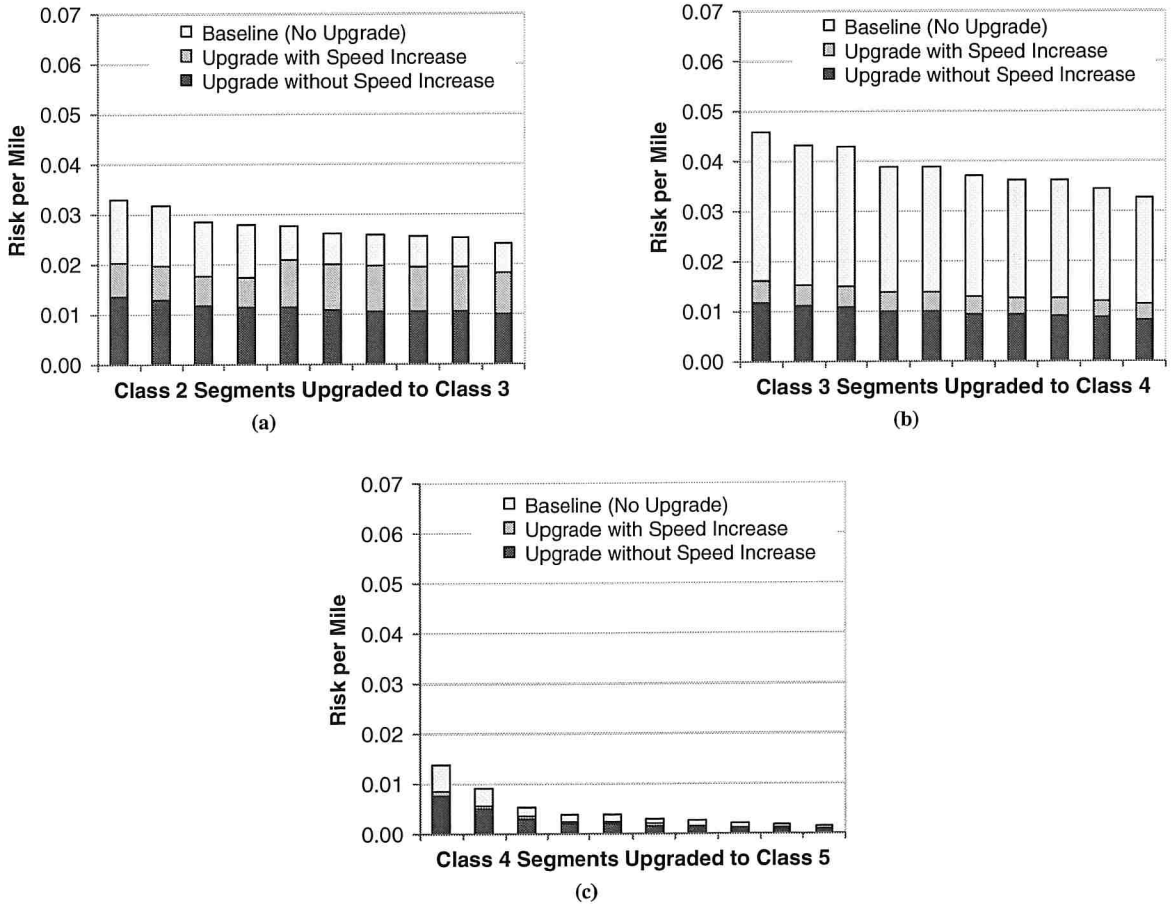


FIGURE 6 Effects of infrastructure upgrade and speed on risk reduction for top 10 segments with highest risk per mile: (*a*) Class 2 upgraded to Class 3, (*b*) Class 3 upgraded to Class 4, and (*c*) Class 4 upgraded to Class 5.

TABLE 2 Summary of Estimates

Track Class Considered for Upgrade to One Higher Class	Total Distance Upgraded (m)	Accident Rate (cars derailed per year)	Release Rate (releases per year)			Annual Risk (persons affected per year)		
				Using Speed-Dependent CPR			Using Speed-Dependent CPR	
			Using Average-Speed CPR	Without Speed Increase	With Speed Increase	Using Average-Speed CPR	Without Speed Increase	With Speed Increase
Before Track Upgrade (baseline)		0.0144	0.0051	0.0066		1.2913	1.4281	
After Track Upgrade								
Class 2	11	0.0140	0.0049	0.0065	0.0066	1.2018	1.3621	1.3881
Class 3	230	0.0093	0.0033	0.0046	0.0050	0.5879	0.6753	0.8003
Class 4	545	0.0125	0.0044	0.0056	0.0059	1.2248	1.3276	1.3484
Percent Reduction from Baseline								
Class 2	11	3.1	3.9	1.5	0.7	6.9	4.6	2.8
Class 3	230	35.5	35.3	30.3	24.2	54.5	52.7	44.0
Class 4	545	13.2	13.7	15.2	10.6	5.1	7.0	5.6
Percent Reduction per Mile Upgraded								
Class 2	11	0.28	0.36	0.13	0.06	0.63	0.42	0.25
Class 3	230	0.15	0.15	0.13	0.11	0.24	0.23	0.19
Class 4	545	0.02	0.03	0.03	0.02	0.01	0.01	0.01

First, in the consideration of track infrastructure upgrades, there are different opportunities for risk reduction based on the distribution of track classes on the route and the differential effects on risk reduction with respect to each track class. Upgrading the lowest track class provides the highest risk reduction per track mile; however, it represents the smallest proportion of the route, so that it does not offer much opportunity to reduce risk. On the other hand, upgrading Class 3 track segments to Class 4 provides the highest risk reduction overall, even though these segments do not represent the largest percentage of the route compared with the percentages for the other track classes.

Second, use of the speed-dependent CPR allows the effects of speed on different routes to be properly accounted for, whereas use of the average-speed CPR does not. Furthermore, the speed-dependent CPR is necessary for proper evaluation of the effect of track-class upgrades on risk with or without a speed increase.

Third, for the representative route, use of the speed-dependent CPR yields annual risk estimates slightly higher than use of the average-speed CPR. This is because the majority of the route has higher-than-average track speeds. This result will vary depending on route-specific characteristics.

Fourth, upgrading the track class with the commensurate speed increase actually reduces the risk relative to that for the baseline case. This is because the incremental effect of the lower accident rate due to the higher track class more than offsets the slight increase in the CPR due to the higher speed.

Completing the analyses used to illustrate the effect of speed-dependent CPR on risk estimates required several assumptions. As mentioned, for track-class upgrades with an increase in speed, it was assumed that trains would operate at the maximum allowable speed corresponding to the upgraded track class. This will result in estimates of speed higher than those that would actually be used on some segments, thus overestimating the risk for those segments. Inferring track class on the basis of the timetable speed may not always represent actual maintenance conditions for some track segments, also resulting in the overestimation of risk. For both of these circumstances, if data for actual operating speed are available, they can be incorporated into the model and use of the speed-dependent CPR will make such adjustments more accurate. Additionally, the relationships developed to estimate the speed-dependent CPR in this study do not account for other possible speed-dependent effects, such as the number of cars derailed or the spill size distribution, but these factors can also be incorporated into the model, if suitable data are available.

The intent of this research was to develop a general method to calculate the speed-dependent CPR by using published data for railroad tank cars. The authors then used a case study to compare several different scenarios that affect or that are affected by speed to gain insight into the effect of using the speed-dependent CPR on estimates of safety and risk. Analysis of these scenarios is not meant to suggest that these are necessarily the most cost-effective approaches to risk reduction. Instead, the methodology presented here is intended to describe and illustrate the utility of speed-dependent CPR and how it may be incorporated into safety and risk analysis calculations. The scenarios considered here include upgrading the track class both with and without speed increases. Either scenario would entail considerable additional capital and maintenance expense (*19*). Rational consideration of these options would require data on their incremental cost and, ideally, information on the cost-effectiveness of other options that might also affect risk. This would enable informed decisions about when and where risk reduction options should be implemented and which ones are the most cost-effective to be made.

CONCLUSIONS

In this paper, the authors present a technique that incorporates the effects of different train speeds on accident rate and CPR and incorporate both of these into a hazardous materials transportation risk

analysis model. The relationship was developed to estimate the speed-dependent CPR, thereby allowing more accurate estimates of route risk. Although previous work has addressed various aspects of this question (*1, 7–9*), the method used to calculate the speed-dependent CPR described here can be used to develop estimates for specific tank car designs, which was not previously possible. Additionally, the statistics and methodology presented here are based on the most up-to-date, published data on tank car performance in accidents. The utility of the speed-dependent CPR is illustrated by using a case study of track infrastructure upgrades in which the authors analyzed the effects of train speed and track improvement independently and together to assess their effects on route risk. Different options in a track infrastructure upgrade problem were considered to determine the option that may provide the greatest reduction in risk for the particular route studied.

Previous work has not explicitly considered the nature or the extent of the effect of using the average-speed CPR versus the speed-dependent CPR on risk estimates. Although use of the average-speed CPR may often be appropriate for national or systemwide estimates of risk, it does not satisfactorily account for differences in operating characteristics when routes are compared. Different routes will almost certainly have different speed distributions and population exposures. Use of the average-speed CPR can lead to both under- and overestimates of the overall risk along the route and at specific locations along the route. Both of these may affect risk management decisions; therefore, the ability to calculate them more accurately is important. In addition, use of the speed-dependent CPR affects the ability to assess the effects of certain types of track infrastructure upgrades quantitatively, as illustrated in the case study. The speed-dependent CPR calculation methodology and its application to risk analysis models should enable more accurate risk calculations not only for comparisons of the risks on different routes but also for assessment and evaluation of the benefits of infrastructure improvements that affect operating speeds.

ACKNOWLEDGMENTS

This research was supported by the Railroad Engineering Program at the University of Illinois at Urbana–Champaign. The authors are grateful to Todd Treichel, Cherry Burke, and the reviewers for their helpful comments and suggestions on this paper.

REFERENCES

1. Barkan, C. P. L., C. T. Dick, and R. Anderson. Railroad Derailment Factors Affecting Hazardous Materials Transportation Risk. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1825,* Transportation Research Board of the National Academies, Washington, D.C., 2003, pp. 64–74.
2. Kawprasert, A., and C. P. L. Barkan. Effects of Route Rationalization on Hazardous Materials Transportation Risk. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2043,* Transportation Research Board of the National Academies, Washington, D.C., 2008, pp. 65–72.
3. Saat, M. R., and C. P. L. Barkan. The Effect of Rerouting and Tank Car Safety Design on the Risk of Rail Transport of Hazardous Materials. *Proc., 7th World Congress on Railway Research,* Montreal, Quebec, Canada, 2006.
4. Glickman, T. S., E. Erkut, and M. S. Zschocke. The Cost and Risk Impacts of Rerouting Railroad Shipments of Hazardous Materials. *Accident Analysis and Prevention,* Vol. 39, 2007, pp. 1015–1025.
5. Verma, M. A Cost and Expected Consequence Approach to Planning and Managing Railroad Transportation of Hazardous Materials. *Transportation Research, Part D,* Vol. 14, 2009, pp. 300–308.
6. Kawprasert, A., and C. P. L. Barkan. Communication and Interpretation of Results of Route Risk Analyses of Hazardous Materials Transportation by Railroad. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2097,* Transportation Research Board of the National Academies, Washington, D.C., 2009, pp. 125–135.
7. Treichel, T. T., J. P. Hughes, C. P. L. Barkan, R. D. Sims, E. A. Phillips, and M. R. Saat. *Safety Performance of Tank Cars in Accidents: Probability of Lading Loss.* Report RA-05-02. RSI-AAR Railroad Tank Car Safety Research and Test Project. Association of American Railroads, Washington, D.C., 2006.
8. Nayak, P. R., D. B. Rosenfield, and J. H. Hagopian. *Event Probabilities and Impact Zones for Hazardous Materials Accidents on Railroads.* Report DOT/FRA/ORD-83/20. FRA, U.S. Department of Transportation, 1983.
9. *Guidelines for Chemical Transportation Risk Analysis.* Center for Chemical Process Safety, American Institute of Chemical Engineers, New York, 1995.
10. Pipeline and Hazardous Materials Safety Administration, U.S. Department of Transportation. *Emergency Response Guidebook.* http://hazmat.dot.gov/pubs/erg/erg2008_eng.pdf. Accessed July 2008.
11. Brown, D. F., H. M. Hartmann, W. A. Freeman, and W. D. Haney. *Development of the Table of Initial Isolation and Protective Action Distances for the 2008 Emergency Response Guidebook.* Publication ANL/DIS-09-2. Argonne National Laboratory, U.S. Department of Energy, 2009.
12. Bureau of Transportation Statistics, Research and Innovative Technology Administration, U.S. Department of Transportation. *National Transportation Atlas Database.* www.bts.gov/publications/national_transportation_atlas_database/. Accessed May 1, 2007.
13. Anderson, R. T., and C. P. L. Barkan. Railroad Accident Rates for Use in Transportation Risk Analysis. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1863,* Transportation Research Board of the National Academies, Washington, D.C., 2004, pp. 88–98.
14. Environmental Systems Research Institute, Inc. (ESRI) Data and Maps. Electronic database, 2005.
15. FRA, U.S. Department of Transportation. *Track Safety Standards.* Vol. 4, 49 CFR 213.9, 2003.
16. *Recommended Railroad Operating Practices for Transportation of Hazardous Materials.* Circular OT-55-J. Association of American Railroads, Washington, D.C., 2009.
17. U.S. Department of Transportation. *Hazardous Materials: Enhancing Rail Transportation Safety and Security for Hazardous Materials Shipments; Railroad Safety Enforcement Procedures; Enforcement, Appeal and Hearing Procedures for Rail Routing Decisions.* Final Rules. 49 CFR Parts 172, 174, and 209. *Federal Register,* Vol. 73, No. 229, 2008.
18. *FRA Guide for Preparing Accident/Incident Reports.* Report DOT/FRA/RRS-22. FRA, U.S. Department of Transportation, 2003.
19. Zarembski, A. M., and R. R. Resor. *Technical Monograph: Estimating Maintenance Costs for Mixed High Speed Passenger and Freight Rail Corridors.* Office of Railroad Development, FRA, U.S. Department of Transportation, 2004.

The Transportation of Hazardous Materials Committee peer-reviewed this paper.

Optimizing Train Network Routing with Heterogeneous Traffic

Yung-Cheng (Rex) Lai, Mark H. Dingler, Chung-En Hsu, and Pei-Chun Chiang

The expected increase in long-term demand for rail traffic in North America will lead to capacity constraints. Although both infrastructure investment and operational changes can relieve congestion, careful consideration of how operational practices can affect or mitigate demand is critical for cost-effective planning for new capacity, especially when the high cost required to build and maintain infrastructure is considered. A critical aspect of capacity management is understanding the effect of heterogeneous traffic on capacity. Greater heterogeneity on a line increases interference between trains and creates more delays than would occur if all trains had similar characteristics. The present research takes into account the impact of heterogeneity and presents a new optimization framework to identify optimal train network routing. This tool will help service design managers with traffic routing and capacity utilization. The computational results show that the use of this tool can successfully reduce operational costs and provide reliable services.

Long-term demand for railroad freight traffic is projected to increase worldwide. AASHTO predicts that freight rail service will increase 84%, on the basis of the number of ton miles, by 2035 (*1*). To accommodate this new traffic, various changes in both operations and infrastructure will be required. Infrastructure expansion is capital intensive, and the need for it is affected by operational practices. Consequently, additional capacity may be achieved by alteration of operations. Efficient planning and investment in capacity projects require understanding of how the operational aspects of new traffic will affect the capacity, as well as how operational changes can mitigate the demand (*2*).

A key factor affecting rail capacity is the interaction of different train types. Heterogeneity in train characteristics causes delays greater than those that a corresponding number of homogeneous trains would. In North America, intermodal, manifest, unit, and local trains may all share trackage. Some lines also have intercity passenger trains and, in metropolitan regions, commuter trains. Each of these train types can have considerably different characteristics, and even trains of the same class may have various weights and lengths. This heterogeneity has a substantial effect on rail line capacity (*3–5*).

Previous work has investigated several of the causes of train type heterogeneity and train delay. In Spain, Abril et al. (*6*) considered trains operating at two speeds, normal and 50% of normal, on single- and double-track lines. In North America, Bronzini and Clarke used a single-track simulation model to compare the delay–volume curves of different mixtures of intermodal and unit trains (*7*). Harrod compared the different impacts of faster and slower nonconforming trains and found that the slower that the nonconforming train was, the greater that the impact on the network was (*8*). Gorman used actual traffic data to estimate delay statistically (*9*). Dingler et al. performed analyses using dispatch simulation software to determine the impacts and causes of heterogeneity with freight and passenger traffic (*5*). They found that heterogeneity increases delay but that different types of heterogeneity had different effects, which has implications for capacity planning and rail operations.

Despite its substantial effect on capacity, heterogeneity has not been taken into account in the train network routing problems in the existing literature (*10–15*). Although railroads do create dedicated routes for certain traffic on specific routes, that process is still largely manual and is based on experience. Consequently, a new decision-support process for optimization of the routing of trains through the network with the consideration of heterogeneity is proposed in this paper. It will help service design managers with traffic routing and capacity utilization.

The following sections first demonstrate the impact of heterogeneity on line capacity and then describe the heterogeneity routing process in detail. These are followed by case studies and conclusions.

IMPACT OF HETEROGENEITY ON DELAY AND CAPACITY

Railroad capacity is influenced by a complex relationship of infrastructure and operational parameters. Operational factors include the average and the variability in operating speeds, traffic volume, stability, terminal efficiency, and the heterogeneity of various train characteristics. These are interrelated with and further influenced by infrastructure characteristics, such as siding length and spacing, crossover spacing, number of tracks, signal and traffic control systems, grade, and curvature. For this work, the focus is the impact of heterogeneity on delay and line capacity.

Rail Traffic Controller dispatching simulation software was used to show the impact of different levels of heterogeneity on capacity on a hypothetical signalized, single-track rail line with the characteristics typical of those in midwestern North America (*16*).

On a single-track line, the effect of additional trains on delay is not linear. Instead, the relationship between train volume and delay is exponential, with each train type and traffic mix having its own particular functional relationship (*7, 17, 18*). By using the specified characteristics, delay–volume curves were developed for unit train, intermodal train, and a combination of the two (Figure 1).

Y.-C. Lai, Room 313, and C.-E. Hsu and P.-C. Chiang, Room 301, Department of Civil Engineering, National Taiwan University, Civil Engineering Building, No. 1, Roosevelt Road, Section 4, Taipei 10617, Taiwan. M. H. Dingler, Railroad Engineering Program, University of Illinois at Urbana–Champaign, B-118 Newmark Civil Engineering Laboratory, 205 North Mathews Avenue, Urbana, IL 61801. Corresponding author: Y.-C. Lai, yclai@ntu.edu.tw.

Transportation Research Record: Journal of the Transportation Research Board, No. 2159, Transportation Research Board of the National Academies, Washington, D.C., 2010, pp. 69–76.
DOI: 10.3141/2159-09

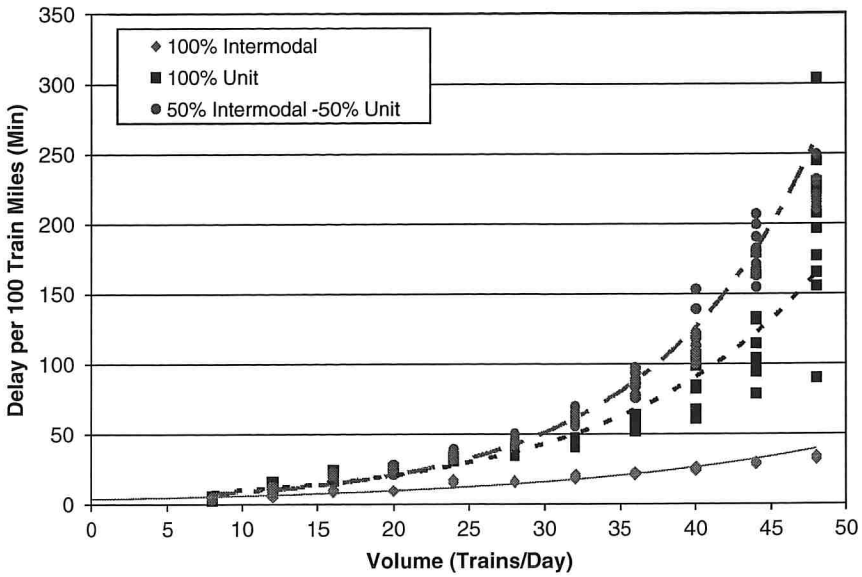


FIGURE 1 Delay per 100 mi versus volume.

When the traffic is heterogeneous, the delay increases at a greater rate to the volume than it does with homogeneous traffic. The interaction of the characteristics of these two trains, that is, the higher speeds, the higher power-to-weight ratios, and the priorities of the intermodal train compared with those of the unit train, creates conflicts that increase the delay over that which would be seen if the traffic was homogeneous with intermodal or unit train traffic.

The results show that the increase in both delay and run time due to heterogeneity are dependent on the train types, volume, and ratio of each train type (*5*). A decrease in heterogeneity could reduce the total run time and could thus reduce the transportation cost. This factor should accordingly be incorporated into the process of decision making over the railway network service design. The next section describes a new traffic routing process, developed in this study, that enables the consideration of heterogeneity.

HETEROGENEOUS TRAFFIC ROUTING PROCESS

Figure 2 shows the process proposed for railroad heterogeneous traffic routing. To calculate the operational cost for train routing, the two most important components are the transportation cost associated with the travel time and the maintenance cost related to tonnage. The proposed process requires the input of the link (subdivision) properties, the run time estimation model, the delay cost, and the traffic demand. Except for traffic demand, the other three inputs are used to develop maintenance cost and transportation cost functions. The output of these functions and traffic demand will then be adopted by the heterogeneous traffic routing module to determine the optimal routing for heterogeneous trains. The following sections demonstrate how the transportation and maintenance costs are determined and how the optimization model is formulated for heterogeneous traffic routing.

Transportation Cost Function

In this study, the transportation cost function was computed on the basis of the product of travel time and the value of travel time. Travel time can be derived from the run time estimate, and the value of travel time is assumed to be the same as the delay cost, which is the sum of the locomotive cost, idling fuel cost, car and equipment cost, and crew cost. A recent estimate of this delay cost by one Class 1 railroad was approximately $261 per train hour (*19*). That value is the default value set in the optimization model; however, different types of trains may have different unit delay costs because of the value and time sensitivity of the loads transported. Therefore, users can input their own data into the heterogeneous traffic routing module.

The travel time for a particular train from its origin to destination is the total run time through all subdivisions on its route. Although the routing of trains is determined by the routing module, estimation of the run time on the basis of the traffic composition for each subdivision is a model input. Railroad managers usually have their own methods or tools that they use to estimate the run times of trains through a subdivision. These methods include statistical models developed on the basis of simulations or historical data (*9, 18, 20*); stringline tools, such as the widely used Intelligent Train Scheduler developed by

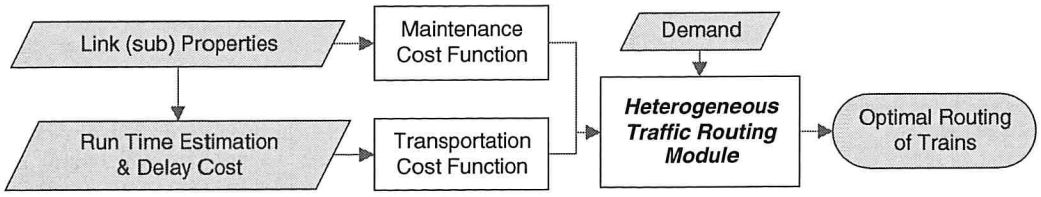


FIGURE 2 Process of optimal heterogeneous traffic routing.

TABLE 1 Coefficient of Regression Models for Subdivision KO (9)

	Free Running Time (FRT)			Congestion-Related Delay (CRD) Variables					
	Train-Related			Primary–Direct Effect			Secondary Effect		
Train Priority	Int.	HPT	Direction	Meet	Overtaken	Pass	Prior (12–24 h)	Prior (6–12 h)	Prior (0–6 h)
High	4.89	(0.0992)	(0.0652)	0.0632	0.5679	0.0127	0.0043	0.0131	0.0385
Medium	5.34	(0.2051)	0.0439	0.1173	0.3355	(0.0062)	0.0017	0.0248	0.0726
Low	5.97	0.0034	0.3533	0.1783	0.6050	(0.0403)	(0.0187)	0.0303	0.1220

NOTE: Int. = intercept; HPT = horsepower per ton.

Intelligent Railroad Systems; and simulation tools, such as the Rail Traffic Controller, the Route Capacity Model, and RAILSIM. The required input for run time for the proposed routing process is the relationship between the run time and the traffic composition for each subdivision, which can be either an estimation model or a data table. The users are responsible for choosing the appropriate model for their application.

Many of the factors affecting the run time (e.g., the track configuration and the traffic mix) can be fixed for a particular subdivision; therefore, the size of the variable set for computation of the run time would be smaller than that for a consolidated run time model that works for all subdivisions. Gorman used historical data to develop a set of linear regression models for each train type in every subdivision that he analyzed (*9*). Table 1 presents the coefficients of important factors of the three regression models estimating the run times of three different types of trains traveling on Subdivision KO (from Dilworth, Minnesota, to Minot, North Dakota). If these run time models for each link in the network are known, they can be used to obtain the required run time data for the proposed routing model. The detailed steps of this process are presented below:

Step 1. Generate all possible scenarios with different traffic compositions (e.g., different amounts of each train type).

Step 2. For each scenario, determine the value of each parameter for the regression model [the number of meets, passes, and overtakes are estimated according to the equation in UIC Leaflet 406 (*4*)]:

Step 2a. Train priority, horsepower per ton, direction, departure headway, and train spacing variability can be obtained from the characteristics of the trains and the base train schedule.

Step 2b. Number of meets = (expected number of trains encountered in the journey in the opposite direction) × (probability of meeting a train with a higher priority + 0.5 × probability of meeting a train with the same priority).

Step 2c. Number of passes = (expected number of trains encountered in the journey in the same direction) × (probability of encountering a train with a lower priority).

Step 2d. Number of overtakes = (expected number of trains encountered in the journey in the same direction) × (probability of encountering a train with a higher priority).

Step 2e. Number of prior trains = expected number of prior trains in 0 to 6 h, 6 to 12 h, and 12 to 24 h on the basis of symmetric traffic dispatching (symmetric in direction and train types).

Step 3. Compute the run time for each type of train by use of the coefficients in Table 1 and parameters in Step 2.

For this work, three train types were attributed to the train priority groupings. Types I, II, and III correspond to trains with high, medium, and low priorities, respectively. Table 2 is an excerpt of the complete run time table for Subdivision KO. There will be a table for each subdivision in the network analyzed as the input for the heterogeneous routing mode.

Maintenance Cost Function

Estimation of the maintenance cost is not an easy task because railroad maintenance costs are not publicized. The best public information can be obtained from an FRA technical monograph: *Estimating Maintenance Costs for Mixed High-Speed Passenger and Freight Rail Corridors* (*21*). That report provides tables estimating the maintenance cost on the basis of tonnage, curvature, track class, and type of ties. For example, Table 3 shows the minimum maintenance cost per track mile for routes with wood ties and predominantly freight trains.

According to the data obtained from the monograph, several matching equations for a particular combination of curvature, type of tie, and track class are then created. For every possible combination, a linear relationship between maintenance cost and tonnage can be established as:

$$C^M = \alpha X + \beta \quad (1)$$

where

C^M = average maintenance cost (dollars/mi),
X = total tonnage [millions of gross tons (MGT)], and
α and β = model coefficients.

Table 4 lists the coefficients for different sets of curvatures, types of ties, and track classes. Curvature is defined by the percentage of

TABLE 2 Excerpt of Estimated Run Time Based on the Traffic Composition of Subdivision KO

Number of Trains			Run Time (h/train)			
Type I	Type II	Type III	Type I	Type II	Type III	Total Run Time (h)
7	20	1	5.73	5.83	6.79	163
3	5	29	5.72	6.21	7.10	254
4	12	12	5.86	5.94	6.85	177
15	3	1	5.25	5.64	6.53	102
26	16	24	7.02	7.44	8.02	494
:	:	:	:	:	:	:

TABLE 3 Annual Minimum Maintenance Cost per Track Mile with Predominant Freight Trains, Light Curves, and Wood Ties

	Tonnage (MGT)			
	≤5 ($)	5–15 ($)	15–30 ($)	≥30 ($)
Class 4	26,554	31,046	42,008	60,833
Class 5	29,172	34,408	46,816	68,778
Class 6	32,962	39,329	53,833	80,265

the total distance with curves of less than 2 degrees, 2 to 6 degrees, or greater than 6 degrees. For example, a line segment with server curvature (82%–16%–2%) means that 82% of the track has curvatures of less than 2 degrees, 16% of the track has curvatures of between 2 and 6 degrees, and 2% of the track has curvatures of greater than 6 degrees. The present application focuses on freight railroads, so track classes higher than Class 5 were ignored. Track Class 1 was also excluded because it is hardly ever used for main-line track. However, the data in the FRA monograph do not provide information on Track Classes 2 and 3 (*21*); therefore, their coefficients (α and β) were extrapolated from the track classes for which data are available (i.e., Track Classes 4 and 5).

For each segment in the network, the specific maintenance cost function is determined on the basis of its curvature, type of tie, and track class; therefore, the only factor left to affect the maintenance cost would be the annual tonnage.

Heterogeneous Traffic Routing Module

The core of the heterogeneous traffic routing module is the optimization model used to determine the optimal routing for heterogeneous traffic. Trains with different origins and destinations can be considered in a manner similar to that used for multiple commodities sharing a common line; therefore, this problem was formulated as a mixed-integer, network design model (*22–24*).

The following notation is used in the optimization model: i is an index referring to the starting node of an arc, and j is the ending node of an arc; t represents the type of train from types 1 to T; k_t corresponds to origin–destination (O-D) pair k_t by train type t traveling from origin o^{k_t} to destination d^{k_t}; D^{k_t} is the demand of O-D pair k_t; V is an index representing the traffic composition, in which each V refers to a particular traffic mix representing a specific combination of number of trains for each type; and v_t^V is the traffic volume per day for type t train under traffic composition V. L_{ij} is the length of link (i, j); RT_{tij}^V is the average run time of type t train on link (i, j) under traffic composition V; C_t^R is the unit value of travel time (in dollars per hour); and C_{tij}^M is the maintenance cost per track mile of type t train on link (i, j).

This model has two sets of decision variables. The first variable is denoted by $x_{ij}^{k_t}$, which is the number of trains running on arc (i, j) from the k_tth O-D pair. The second variable is a binary variable used to determine the traffic composition (V) of link (i, j), denoted by H_{ij}^V:

$$H_{ij}^V = \begin{cases} 1 & \text{if traffic composition } V \text{ is assigned to arc } (i,j) \\ 0 & \text{otherwise} \end{cases}$$

The optimization model is formulated as follows:

$$\min \sum_i \sum_j \sum_V \sum_t v_t^V H_{ij}^V \left(RT_{tij}^V C_t^R + L_{ij} C_{tij}^M \right) \tag{2}$$

subject to

$$\sum_V H_{ij}^V = 1 \ \forall (i,j) \qquad i \neq j \tag{3}$$

$$\sum_{k_t} x_{ij}^{k_t} \leq \sum_V v_t^V H_{ij}^V \qquad \forall (i,j), i \neq j, t \tag{4}$$

$$\sum_j x_{ij}^{k_t} - \sum_j x_{ji}^{k_t} = \begin{cases} D^{k_t} & \text{if } i \in o^{k_t} \\ -D^{k_t} & \text{if } i \in d^{k_t} \\ 0 & \text{otherwise} \end{cases} \qquad \forall i, k_t$$

$$x_{ij}^{k_t} \geq 0 \qquad x_{ij}^{k_t} \in \text{integer}, H_{ij}^V \in \{0,1\} \tag{5}$$

The objective function is to minimize the sum of the transportation cost and the maintenance cost. $RT_{tij}^V C_t^R$ is the transportation cost per train for type t train transporting on arc (i, j) with traffic composition V. As mentioned before, it is based on the running time for train type t on arc (i, j) under a specific traffic composition, V. For each subdivision, the unit maintenance cost (C_{tij}^M) is fixed for a particular type of train according to its typical tonnage. Hence, $\Sigma_t v_t^V H_{ij}^V C_{tij}^M$ is the total maintenance cost of link (i, j) under traffic composition V. Equation 3 ensures that there can be at most one selected traffic composition for every arc (i, j). Equation 4 ensures that the volume of each link is consistent in either H_{ij}^V or Σx_{ij}^{kt}. Finally, Equation 5 is

TABLE 4 Coefficients for Maintenance Cost Functions

Tie	Curvature		Class 2	Class 3	Class 4	Class 5
Wood	Light (96%–4%–0%)	α	474.2	619.4	773.3	892.2
		β	17,082.0	19,196.0	20,777.0	24,489.0
	Moderate (90%–8%–2%)	α	499.3	651.0	811.7	935.9
		β	16,574.0	18,784.0	21,134.0	22,925.0
	Severe (82%–12%–6%)	α	542.9	703.6	873.9	1,005.6
		β	16,813.0	19,155.0	21,644.0	23,545.0
Concrete	Light (96%–4%–0%)	α	406.8	550.9	703.7	821.6
		β	17,182.0	19,279.5	21,510.0	23,208.0
	Moderate (90%–8%–2%)	α	423.7	575.3	736.9	858.7
		β	17,100.0	19,361.0	21,790.0	23,546.0
	Severe (82%–12%–6%)	α	471.9	627.2	791.8	919.0
		β	17,526.0	19,789.0	22,195.0	24,029.0

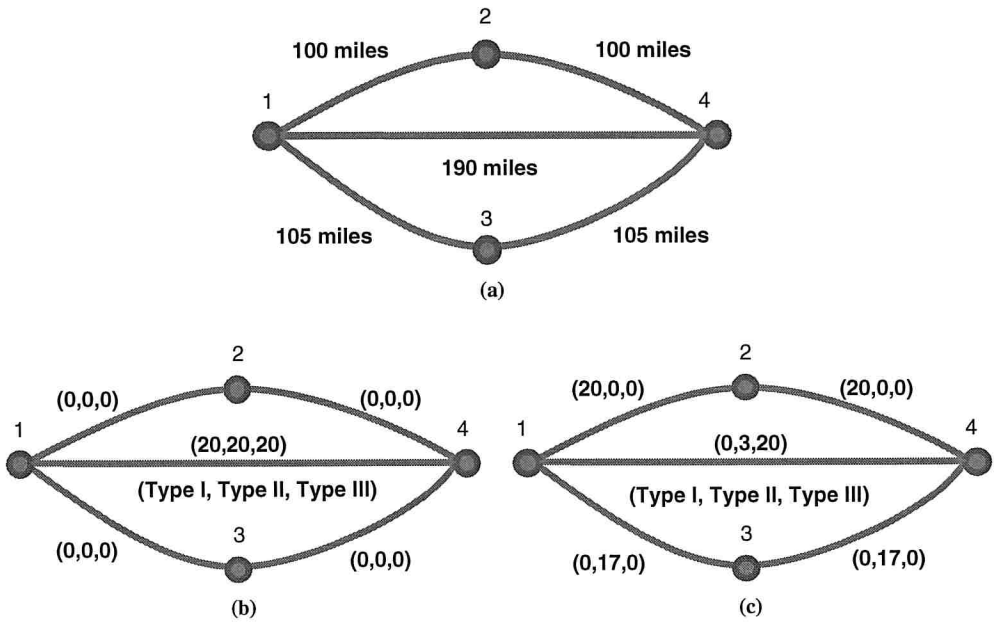


FIGURE 3 Case Study I (*a*) network, (*b*) shortest-path routing, and (*c*) optimal routing.

the network flow conservation constraint guaranteeing that the outflow is always equal to the inflow for transshipment nodes; otherwise, the difference between them should be equal to the demand for that O-D pair.

CASE STUDY

To demonstrate the potential use of this tool, the proposed process was applied to two case studies and the computational results from the optimal network routing and from the shortest-path routing were compared. The shortest-path routing may not exactly describe the real-world routing; however, it can be considered the upper bound of the problem. Consequently, the range of the routing solutions can be determined with this upper bound and the lower bound obtained from the optimal solution. The size of the network in Case Study I was intended to be small to easily show the difference in routing obtained by different strategies and the potential benefit of using this tool. Case Study II is a more practical case with higher traffic demand and a larger network (about 4,000 mi of track) to prove the applicability of this model to a real problem.

Case Study I

Figure 3*a* shows the network for Case Study I. It is assumed that for each train type there are 10 trains to be dispatched from Node 1 to Node 4 and another 10 trains to be dispatched from Node 4 to Node 1. For every train, one of three different routes can be chosen: (*a*) Point 1 to Point 4 (190 mi), (*b*) Point 1 to Point 2 to Point 4 (200 mi), and (*c*) Point 1 to Point 3 to Point 4 (210 mi). Each link in the network was assumed to be similar to Subdivision KO, mentioned above, except for its length, so the same running time estimation method (Table 2) and maintenance cost function were used. This case study was coded in General Algebraic Modeling System (GAMS) software and solved by use of the CPLEX optimization solver within minutes with a 3-GHz computer (*25*).

Compared with the optimal routing, the shortest-path algorithm did not take into account the impact of heterogeneity but allocated all trains through the path from Point 1 to Point 4 (Figure 3*b*). As a result, the path from Point 1 to Point 4 became highly congested, so the transportation cost increased substantially (Table 5). Figure 3*c* represents the optimal routing obtained from the heterogeneous traffic routing module. Different types of trains were dispatched through different paths; however, three Type II trains took the path from Point 1 to Point 4 with the majority of Type III traffic because the shorter routes compensated for the impact on heterogeneity. In this case, the total run time under the shortest-path algorithm was about 5.97 h per train per day, whereas the heterogeneous routing model effectively reduced the total run time to 4.88 h per train per day. Consequently, the total cost of the optimal routing is about 10% lower than that of shortest-path routing (Table 5).

Case Study II

In Case Study II, the heterogeneous traffic routing process was applied to a network with 25 nodes and 40 links (Figure 4). The size of the network is about 4,000 mi of track, and the symmetric traffic demand for every O-D pair is listed in Table 6. Similar to Case Study I, each link in the network was assumed to be similar to Subdivision KO, so the same running time estimation (Table 3) and maintenance cost function were used. The case was designed to be a close representation of the real-world problem. To easily compute the heterogeneity

TABLE 5 Computational Results Based on Different Routing Strategies in Case Study I

Routing Study	Maintenance Cost ($)	Transportation Cost ($)	Total Cost ($)
Shortest path	23,473,040	34,148,120	57,621,160
Optimum	24,154,440	27,945,800	52,100,230

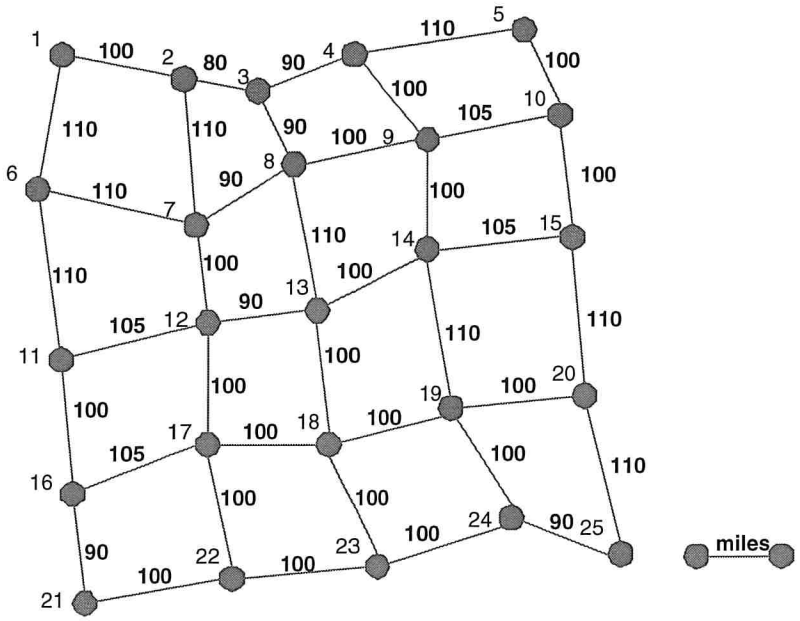


FIGURE 4 Network of Case Study II.

of each link, only Type I traffic and Type III traffic were used in this case, so the heterogeneity of each link can be defined as follows:

$$HG_{ij} = \min\left(\frac{N_{ij}^{t}}{N_{ij}^{\text{total}}}, 1-\frac{N_{ij}^{t}}{N_{ij}^{\text{total}}}\right)\times 100\% \tag{6}$$

where

HG_{ij} = heterogeneity of link (i, j),
N_{ij}^{t} = number of type t trains on link (i, j), and
N_{ij}^{total} = total number of trains on link (i, j).

TABLE 6 Traffic Demand of Case Study II

Type	Origin	Destination	Demand (trains/day)
I	1	25	3
III	1	25	3
I	25	1	5
III	25	1	5
I	21	5	5
III	21	5	5
I	5	21	3
III	5	21	3
I	6	20	5
III	6	20	5
I	20	6	5
III	20	6	5
I	16	10	3
III	16	10	3
I	10	16	3
III	10	16	3

HG_{ij} equal to 50% represents the highest degree of heterogeneity, and HG_{ij} equal to 0% represents homogeneous traffic.

Similar to Case Study I, Case Study II was also coded in GAMS software and solved by the use of the CPLEX optimization solver (*25*); however, the size of the problem increased the solution time to about 3 h. Both the shortest-path routing and the optimal routing are demonstrated in Figure 5, and Table 7 summarizes the computational results.

The network heterogeneity was calculated on the basis of the average heterogeneity of every link in the network. In this example, the optimal routing significantly reduced its heterogeneity compared with that of the shortest-path scenario (Table 7). In addition, the optimal routing required a slightly higher maintenance cost but a much lower transportation cost than the shortest-path routing. As a result, the annual total cost was reduced by \$7.6 million, a 3.41% savings.

DISCUSSION OF RESULTS

The operational cost benefit of using the routing process proposed here has been demonstrated in the case studies. In addition to cost savings, the heterogeneous routing process can also increase the reliability of the system by evenly distributing traffic throughout the network. This is done automatically by the running time model because the delay for heterogeneous traffic increases at a greater rate to volume than that for homogeneous traffic. Moreover, the routing module avoids breakdown by keeping the traffic volume at an appropriate level (Figure 1); otherwise, the increase in the run time and the transportation cost will be undesirable. Within the appropriate level, the system will be better able to recover from an emergency to increase its reliability.

This study focuses on the effect of the traffic mix, so the effect of directional running was not emphasized. However, the proposed heterogeneity routing process is general enough to be able to take into account the impact of directional running. For every link in the network, if the same type of train (e.g., unit trains) running in dif-

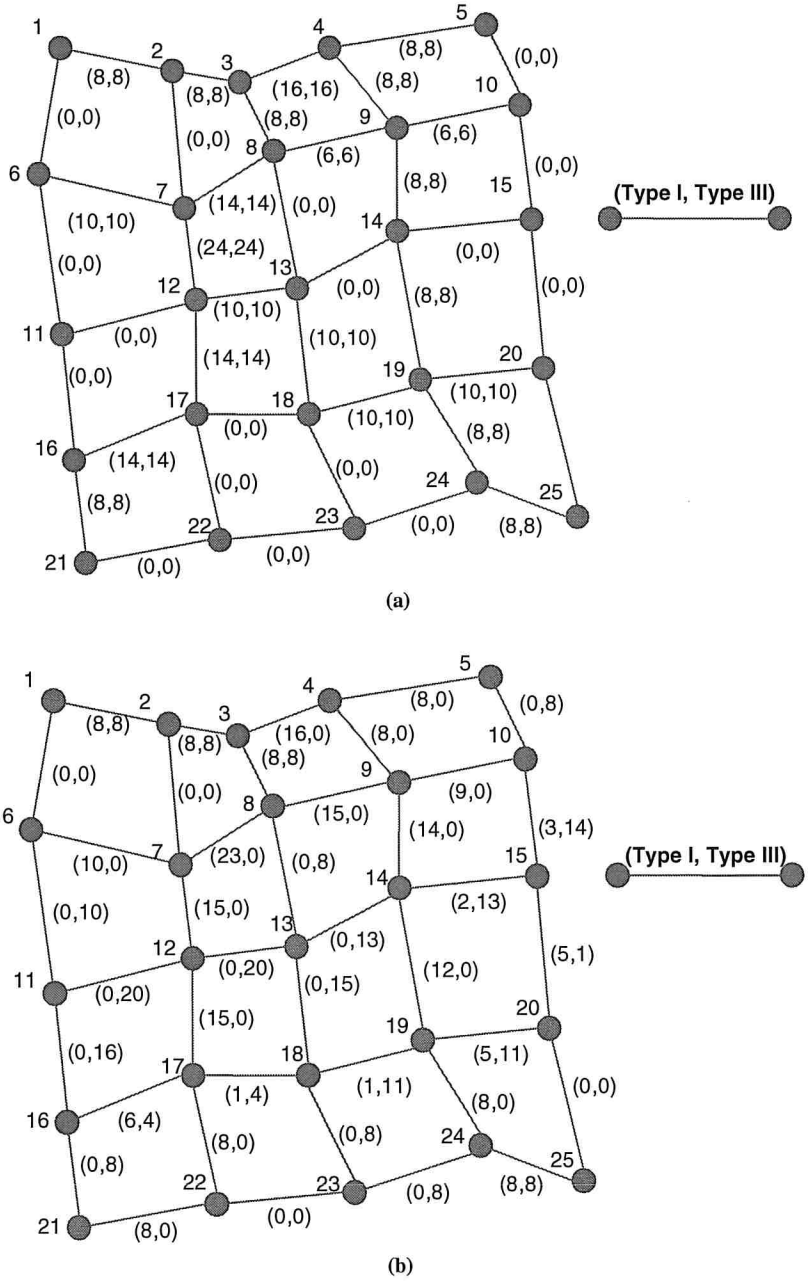


FIGURE 5 Case Study II (*a*) shortest-path routing and (*b*) optimal routing.

TABLE 7 Computational Results Based on Different Routing Strategies in Case Study II

Routing Strategy	Network Heterogeneity (%)	Maintenance Cost ($)	Transportation Cost ($)	Total Cost ($)
Shortest path	27	111,999,692	111,246,799	223,246,491
Optimum	8	112,783,057	102,859,288	215,642,345

ferent directions is considered to be different types (e.g., eastbound unit trains and westbound unit trains), the proposed heterogeneous routing module will be able to take into account the impact from heterogeneity and directional running simultaneously. Consequently, this routing module is highly beneficial for service design managers with regard to traffic routing as well as capacity utilization.

CONCLUSIONS

Long-term demand for worldwide railroad freight traffic is projected to increase. To accommodate this new traffic, various changes in both operations and infrastructure are inevitable. The efficient use of capacity requires an understanding of how the mixture of traffic affects capacity. The research presented here provides a new optimization framework for the identification of the optimal train network routing with the consideration of heterogeneity. The computational results show that the proposed routing process has the potential to reduce operational costs and increase system reliability.

ACKNOWLEDGMENTS

The authors are grateful to Eric Wilson and Christopher P. L. Barkan for their assistance with this research. Support for the second author's graduate study has been from a CN Railroad Engineering Research Fellowship at the University of Illinois at Urbana–Champaign.

REFERENCES

1. *Transportation—Invest in Our Future: America's Freight Challenge.* AASHTO, Washington, D.C., 2007.
2. Lai, Y.-C., and C. P. L. Barkan. Enhanced Parametric Railway Capacity Evaluation Tool. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2117,* Transportation Research Board of the National Academies, Washington, D.C., 2009, pp. 33–40.
3. Pachl, J. *Railroad Operating and Control.* VTD Rail Publishing, Mountlake Terrace, Wash., 2002.
4. UIC Leaflet 406. International Union of Railways, Paris, 2005.
5. Dingler, M. H., Y.-C. Lai, and C. P. L. Barkan. Impact of Train Type Heterogeneity on Single-Track Railway Capacity. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2117,* Transportation Research Board of the National Academies, Washington, D.C., 2009, pp. 41–49.
6. Abril, M., F. Barber, L. Ingolotti, M. A. Salido, P. Tormos, and A. Lova. An Assessment of Railway Capacity. *Transportation Research, Part E: Logistics and Transportation Review,* Vol. 44, No. 5, 2008, pp. 774–806.
7. Bronzini, M. S., and D. B. Clarke. Estimating Rail Line Capacity and Delay by Computer Simulation. *Transportation Forum,* Vol. 2, No. 1, 1985, pp. 5–11.
8. Harrod, S. Capacity Factors of a Mixed Speed Railway Network. *Transportation Research, Part E: Logistics and Transportation Review,* Vol. 45, 2009, pp. 830–841.
9. Gorman, M. F. Statistical Estimation of Railroad Congestion Delay. *Transportation Research, Part E: Logistics and Transportation Review,* Vol. 45, 2009, pp. 446–456.
10. Crainic, T. G., J. A. Ferland, and J. M. Rousseau. A Tactical Planning Model for Rail Freight Transportation. *Transportation Science,* Vol. 18, 1984, pp. 165–184.
11. Haghani, A. E. Formulation and Solution of a Combined Train Routing and Makeup, and Empty Car Distribution Model. *Transportation Research, Part B: Methodological,* Vol. 23, No. 6, 1989, pp. 433–452.
12. Keaton, M. H. Designing Optimal Railroad Operating Plans: Lagrangian Relaxation and Heuristic Approaches. *Transportation Research, Part B: Methodological,* Vol. 23, No. 6, 1989, pp. 415–431.
13. Martinelli, D. R., and H. Teng. Optimization of Railway Operations Using Neural Networks. *Transportation Research, Part C: Emerging Technologies,* Vol. 4, 1996, pp. 33–49.
14. Cordeau, J. F. A Survey of Optimization Models for Train Routing and Scheduling. *Transportation Science,* Vol. 32, No. 4, 1998, pp. 380–404.
15. Li, K. P., Z. Y. Gao, B. H. Mao, and C. X. Cao. Optimizing Train Network Routing Using Deterministic Search. *Networks and Spatial Economics,* 2009. DOI: 10.1007/s11067-009-9098-7.
16. Washington Group International, Inc. *RTC Simulations—LOSSAN North Railroad Capacity and Performance Analysis.* LOSSAN Rail Corridor Agency, San Diego, Calif., 2007. www.sbcag.org/Meetings/SCSPC/2007/February/Item5LOSSANSR.pdf. Accessed March 15, 2009.
17. Reistrup, P. H. Railroad Capacity and Corridor Planning. Presented at 81st Annual Meeting of the Transportation Research Board, Washington, D.C., 2002.
18. Krueger, H. Parametric Modeling in Rail Capacity Planning. *Proc., Winter Simulation Conference,* Phoenix, Ariz., 1999.
19. Lai, Y. C. *Increasing Railway Efficiency and Capacity Through Improved Operations, Control and Planning.* PhD dissertation. Department of Civil and Environmental Engineering, University of Illinois at Urbana—Champaign, 2008.
20. Huisman, T., and R. J. Boucherie. Running Times on Railway Sections with Heterogeneous Train Traffic. *Transportation Research, Part B: Methodological,* Vol. 35, 2001, pp. 271–292.
21. Zarembski, A. M., R. R. Resor, J. F. Cikota, R. U. Cogswell, and D. Plotkin. *Estimating Maintenance Costs for Mixed High-Speed Passenger and Freight Rail Corridors.* Technical Monograph. Unpublished report. FRA, U.S. Department of Transportation, 2004.
22. Magnanti, T. L., and R. T. Wong. Network Design and Transportation Planning: Models and Algorithms. *Transportation Science,* Vol. 18, No. 1, 1984, pp. 1–55.
23. Minoux, M. Network Synthesis and Optimum Network Design Problems: Models, Solution Methods, and Applications, *Networks,* Vol. 19, 1989, pp. 313–360.
24. Ahuja, R. K., T. L. Magnanti, and J. B. Orlin. *Network Flows: Theory, Algorithms, and Applications.* Prentice-Hall, Englewood Cliffs, N.J., 1993.
25. *GAMS—A User's Guide.* GAMS Development Corporation, Washington, D.C., 2008.

The Railroad Operating Technologies Committee peer-reviewed this paper.

Effects of Communications-Based Train Control and Electronically Controlled Pneumatic Brakes on Railroad Capacity

Mark H. Dingler, Yung-Cheng (Rex) Lai, and Christopher P. L. Barkan

Railroads are increasingly using new technologies to improve capacity and operating efficiency. To plan their investments and prepare for the implementation of these technologies, railroads must understand their net effect on operations. This requires understanding both the particular aspects of these technologies that affect capacity and the characteristics of the systems into which they are being introduced. Two important technologies in this regard are communications-based train control (CBTC) and electronically controlled pneumatic (ECP) brakes. Each element of CBTC and ECP brakes with the potential to affect capacity was identified, and its effect under various implementation scenarios was evaluated. The potential impact of each element was assessed and compared with the various baseline conditions and conventional technologies to understand the incremental effect. An extensive review of the literature on the subject was conducted in support of these evaluations. CBTC implementation with enforcement braking will generally result in a loss of capacity, but as these systems become more fully integrated, the potential for capacity enhancement improves. ECP brakes will provide benefits under most operational scenarios because of the shorter braking distances and thus the potential for the closer spacing of trains. The two technologies have a potential interactive effect: CBTC may make it possible to more effectively take advantage of one of the principal benefits of ECP brakes: shorter stopping distances. The results for either technology will be route and network specific, so individual railroads will need to conduct analyses to understand the net effect on the capacity of their systems.

Beginning in the early 2000s, major North American railroads were increasingly experiencing capacity constraints, and long-term projections indicate substantial further growth in freight traffic (*1, 2*). Furthermore, new initiatives to expand intercity passenger rail operations on freight railroads will have a disproportionate impact on capacity because of the differences in operational characteristics between freight and passenger trains (*3–6*). Consequently, understanding of the factors that affect rail capacity and the options available to improve it cost-effectively is important.

Infrastructure expansion will undoubtedly play an important role in accommodating new traffic demand; however, two new technologies that will also affect rail capacity are being introduced: communications-based train control (CBTC) [often referred to as positive train control (PTC) in the United States] and electronically controlled pneumatic (ECP) brakes. Both offer safety benefits, and both have been touted as offering capacity benefits as well, but in actuality, the situation is more complicated. These technologies can enhance capacity under some circumstances, have little or no effect under others, and in some cases may actually reduce capacity. Consequently, understanding their net effect on a particular rail line or network requires understanding the status quo of the system into which they are being introduced and in what manner they are being introduced. This paper attempts to identify each critical aspect of these technologies that has the potential to affect capacity and consider what this effect will be under which implementation conditions. Because both of these systems require significant investment from the railroads [estimates range up to $10 billion for PTC (*7*) and over $6.5 billion for full ECP brake implementation (*8*)], if the capacity impacts of these two technologies can be better understood, railroads can make more informed decisions about their implementation.

The work described here is part of a larger effort in which simulation analyses are being conducted and mathematical models are being developed to quantify the effects of these technologies on rail capacity. Railroads planning for the implementation of CBTC and ECP brakes will need to conduct similar assessments to understand the particular effect on their own networks.

CBTC is a system in which train monitoring and train control are integrated into a single system via data links between vehicles, central office computers, and wayside computers (IEEE Standard 1474.2-2003). ECP brakes use an electronic signal instead of the train-line air pressure to transmit braking signals. CBTC has been under development since the mid-1980s (*9–11*), and freight railroad ECP brake technology has been under development since the early 1990s (*12*); however, wide-scale adoption has not occurred because of technical, practical, economic, and institutional barriers (*13*). Recent regulations and legislation have altered the situation. FRA is encouraging the implementation of ECP brakes by offering relief from certain requirements pertaining to conventional pneumatic brake operation (*14, 15*). With regard to PTC, the Rail Safety Improvement Act of 2008 (H.R. 2095, 110th Cong., 2nd session, 2008) and the subsequent regulations issued by FRA (*16*) have mandated its implementation on a large portion of the main lines of Class 1 railroads by 2015.

M. H. Dingler, Railroad Engineering Program, Department of Civil and Environmental Engineering, B-118 Newmark Civil Engineering Laboratory and C. P. L. Barkan, Department of Civil and Environmental Engineering, 1203 Newmark Civil Engineering Laboratory, University of Illinois at Urbana–Champaign, 205 North Mathews Avenue, Urbana, IL 61801. Y.-C. Lai, Department of Civil Engineering, National Taiwan University, Room 313, Civil Engineering Building, No. 1, Roosevelt Road, Section 4, Taipei 10617, Taiwan. Corresponding author: M. H. Dingler, dingler2@illinois.edu.

Transportation Research Record: Journal of the Transportation Research Board, No. 2159, Transportation Research Board of the National Academies, Washington, D.C., 2010, pp. 77–84.
DOI: 10.3141/2159-10

A number of previous studies have investigated the impact of CBTC on capacity. Lee et al. determined that moving blocks could increase the capacity of the Korean high-speed railway (*17*). Another study quantified the capacity benefits of the European Train Control System (ETCS), Europe's version of CBTC (*18*). In the United States, Smith and colleagues studied the potential benefits of Burlington Northern's Advanced Railroad Electronics System and other possible CBTC systems (*19–22*). They calculated how the more efficient meet–pass planning and the increased dispatching effectiveness possible with CBTC will affect capacity. Martland and Smith calculated the potential terminal efficiency improvements resulting from the estimated increases in reliability offered by CBTC (*23*). Although many authors have claimed that a CBTC system with moving blocks will increase capacity (*7, 10, 22–27*), there has been some debate about whether this will in fact be the case (*13, 28*).

Less work has addressed the capacity effects of ECP brakes. Most agree that they will reduce stopping distances and, when they are fully implemented, will allow closer spacing of trains; however, the incremental effect of this reduction will be affected by what other technologies are already in use. Furthermore, taking advantage of this will often require changes in the signal system.

As discussed above, the effects of CBTC and ECP brakes will be context specific; that is, under some circumstances one or both technologies, either alone or in combination, will have the potential to increase capacity, in other cases they will have little or no effect, and in some cases they may reduce capacity. Consequently, the net effect of these technologies on capacity will be determined by the magnitude of these context-specific impacts and the relative frequency at which they occur over a particular route or network.

ELEMENTS OF A CBTC SYSTEM THAT WILL AFFECT CAPACITY

In North America most of the potential CBTC systems are still under development. Although the specific technical details remain unclear, in general, each will have similar features and capabilities. These systems are characterized by the data links that provide better information to dispatchers and train crews. This has the potential to increase efficiency though better train management and control (*29*). However, to comply with the legislative requirements for PTC, they must also prevent train-to-train collisions, overspeed derailments, incursions into established work zone limits, and the movement of a train through a switch left in the wrong position (*16;* Rail Safety Improvement Act of 2008, H. R. 2095, 110th Cong., 2nd session, 2008). The legislation is a performance standard and does not specify the technology that must be used to meet the requirements. In principle, CBTC can be implemented without enforcement braking; however, this has been envisioned as an element of CBTC since the earliest concepts of its development (*9*). It is also technically possible to meet the PTC requirements without the use of a pure CBTC system (*30*); however, most PTC systems in the United States will likely be some form of pure or hybrid CBTC system with enforcement. Because they are not part of the PTC regulation, the additional elements available with a CBTC system will not necessarily be part of a PTC-compliant system, and therefore, the potential benefits or costs of PTC and CBTC are different. This paper considers the potential elements of a CBTC system that may affect capacity, including those required to meet the PTC requirements.

CURRENT TRAFFIC CONTROL SYSTEMS

Most current automatic traffic control systems use wayside signals to manage train speed and headway. Signal spacing is typically set on the basis of the distance that it takes for the worst-case train that normally operates on a line using normal service braking to stop from the maximum permitted train speed at a location. Because the signals are designed for this worst-case train, many trains may have stopping distances shorter than the signal system of the line was engineered for. Furthermore, although the rules of individual railroads vary on the exact language, an engineer is usually required to begin reducing the speed when the train passes a signal displaying a restrictive signal. This means that for a train to continuously maintain normal track speed, it must not encounter signals less favorable than "clear." Consequently, trains must generally be separated by at least two blocks in a three-aspect system and three blocks in a four-aspect system. Because of these operating rules and the use of worst-case braking distances, trains are separated by a distance several times longer than their braking distance.

A variety of traffic control systems are currently in use on North American railroads, but they can be broadly categorized into two types: (*a*) those in which a manual system of spoken or written messages conveys movement authority to trains and (*b*) those in which the dispatcher conveys this authority directly via the wayside signals. Track warrants are the most commonly used manual system, especially on lower-density lines. The capacity on these lines can be increased by overlaying them with automatic block signals, but the authority is still conveyed manually. If more capacity is needed, it can be upgraded to centralized traffic control (CTC), in which the signals themselves convey movement authority. On some track warrant systems and all CTC systems, the dispatcher is able to remotely control switches, allowing the more efficient planning and management of meets and passes of multiple trains on a line.

There are technologies that offer further improvement in operational efficiency, and some of these provide more information to train crews and others that help dispatchers. The oldest of these is cab signals, which take advantage of the coded track circuits in the rails that communicate the aspect information to the wayside signals. Specialized equipment on the locomotive enables the current signal block aspect to be displayed in the cab. With wayside signals, a signal ahead may change to a more favorable indication but the locomotive engineer does not know this until the next signal comes into view. Cab signals allow the engineer to know immediately if a more favorable indication applies and to take advantage of it immediately. Another technology that assists the dispatcher in managing all the traffic on a line is computer-aided dispatching (CAD). In these systems the computer accounts for the operational characteristics of trains and the features of a route to help the dispatcher better plan meets and passes.

ELEMENTS OF A CBTC SYSTEM

A PTC-compliant CBTC system has several components and features that have the potential to affect capacity, either positively or negatively. These are

- Enforcement braking,
- Real-time train operating and location data,
- In-cab displays, and
- Flexible moving blocks.

Enforcement braking is necessary to comply with PTC requirements. Real-time train operating and location data give the dispatcher additional information. This additional information can also be provided to the locomotive on an in-cab display. CBTC also potentially permits the use of flexible moving blocks. Each of these components will affect railroad operations and capacity and will be considered separately.

Enforcement Braking

The element of a PTC system mandated by regulation is enforced braking to prevent unsafe situations. The intent is that the system will stop the train automatically if the engineer fails to take appropriate action to prevent the train from violating its authority limits or speed restrictions. To provide continuous enforcement, an onboard computer must determine when a train must begin braking. This computed braking curve is composed of the distances traveled during (*31*)

- Equipment reaction time,
- Propulsion removal,
- Brake buildup, and
- Full-service brake application.

These distances are highly dependent on various factors, including the initial speed of the train, the train length, car weights, braking efficiency, operative brakes, the brake propagation rate, adhesion, and rail condition. These factors are not accurately known when a train leaves the terminal, resulting in considerable uncertainty in the exact braking distance required (*13, 32*) (Figure 1). For safe operations, a train must have close to a zero probability of an overshoot [FRA has targeted 0.000005, or 5 chances in a million (*13, 33*)]. This necessitates a conservative braking algorithm that considers the worst-case condition for each of the unknown variables. This causes the enforced braking distance to be greater than the average braking distance (*31*). Consequently, the brake application with a PTC system will begin earlier than required for a typical full-service brake application. With or without braking enforcement, a train will brake in the same distance; consequently, an earlier application will cause the train to stop sooner than the engineer intends (*33*). Simulations have shown that the difference between the average stopping distance and the enforced target can be greater than 1,700 ft (*33*).

Braking enforcement can have several negative effects on capacity, including the following:

- An unacceptably large number of trains are forced to start slowing much earlier than normal service braking to prevent enforcement from taking over, slowing the overall operation.
- Train crews are not able to prevent enforcement and thus stop well short of the target.
- Train crews experience difficulty closely approaching a target stopping point, such as when they are pulling into a siding, potentially causing the back of the train to remain on the main line, blocking traffic (*13*).

Work is under way to create more accurate and adaptive braking algorithms (*13*). However, trains may travel long distances after they depart a terminal without making enough brake applications to obtain data adequate for the development of sufficiently accurate, updated estimations of braking distance (*33*), and there will always be some difference between the calculated braking distance and the actual or performance braking distance (*31*). The magnitude of this difference is dependent on the conservativeness of the braking algorithm used; a more conservative algorithm will increase the difference between the actual and the enforcement braking distances. The probability of overshoot used is dependent on the current specifications regarding enforcement braking; consequently, the manner in which those specifications are interpreted will have a direct impact on the effect of enforcement braking on capacity.

It is also possible that enforcement may have little or no impact on operations or capacity. Current wayside signal spacing is based on the braking distance of the worst-case train plus an additional margin of safety. Signal spacing may be greater than the enforced braking distance; therefore, if signals are still used, trains will begin to slow down in response to them instead of the enforcement. Additionally, enforcement algorithms are based on a full-service brake application. In most cases the engineer makes use of dynamic brakes and slows the train at a more gradual rate than with a full-service brake application, potentially preventing enforcement.

Depending on the railroad's operations and rules, enforcement braking has the potential to either increase travel times for the affected train or have no impact at all. If trains are slowed, they may also delay following trains, further reducing capacity. Further discussions and explanations of braking enforcement, adaptive braking, and their implications can be found in papers by Thurston (*31*) and Moore Ede et al. (*13*).

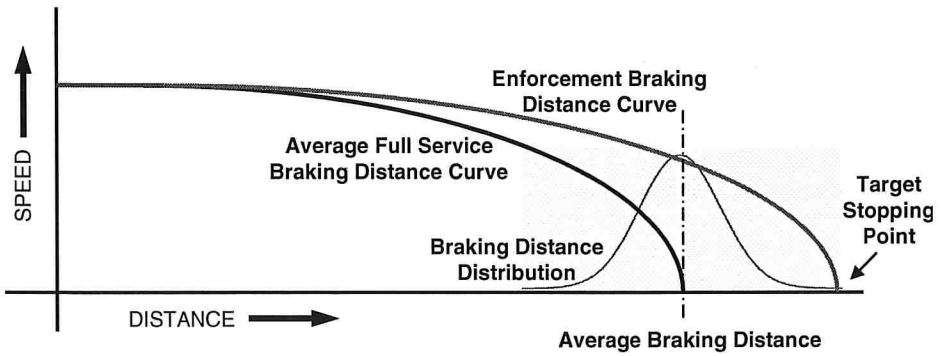


FIGURE 1 Speed, uncertainty in average braking distance, and resultant safe braking distance.

Real-Time Train and Location Data

Real-time train and location data offer the dispatcher additional information. The dispatcher is able to know a train's location and current speed with accuracy and with more precision than existing train control systems provide. This information will allow train dispatchers to respond more quickly to any disruptions or changes and to formulate alternative dispatching plans more quickly as circumstances change. This information also permits more effective meet–pass planning. When these data are combined with a CAD system, they can potentially decrease run times by reducing the time that trains wait for meets and passes (*19, 21*).

Real-time train and location data are also vital to braking enforcement and moving blocks. A technical challenge that has been encountered with real-time data is communications delay in the data links. In a CBTC system, a train's movement depends on the receipt of periodic authority updates as the track ahead clears. Any limitations in the data link throughput and message reliability could limit train capacity. If the data link delivers a movement authority too late, the train may have to reduce speed. Unreliability in the system could result in train position information being inaccurate, to the extent that the uncertainty buffer distances must be increased, increasing train headways (*33*). If the communications delay is not excessive, real-time train and location data can increase capacity.

In-Cab Displays

In-cab displays offer additional information to the locomotive engineer, permitting the engineer to operate the train more efficiently. An in-cab display will most likely have the following information (*34*):

- Location information,
- Authority and speed limits,
- Route and route integrity,
- Start of warning and enforcement braking,
- Location of maintenance-of-way work limits, and
- Positions of other track vehicles.

An in-cab display offers the engineer near real-time information on the status of blocks ahead. With wayside signals, this information is updated only at discrete points as the train approaches and passes each block signal. If the signal is anything less favorable than clear, the engineer will immediately or soon need to reduce speed, unless the train is already traveling at the speed indicated by the signal. Although the status of the block ahead may improve after the front of the train has passed, the engineer has no way of knowing this and will continue reducing the speed until the next signal comes into view and is displaying a more favorable indication. However, if the engineer has access to continuously updated information on the status of the block ahead, the engineer may not have to reduce speed as much if the block ahead clears. A CBTC in-cab display can also have benefits in territories where movement authority is given through a manual system because it eliminates the time required for voice transmission and confirmation (*24*). Cab signal technology provides some of the capacity benefits of a CBTC in-cab display by displaying the aspect of the next block (*31*); however, most locomotives and routes in North America are not equipped with these technologies, so in these cases, CBTC will provide these incremental benefits.

Flexible Moving Blocks

Moving blocks provide continuous train separation and have the potential for this to be based on each train's individual stopping characteristics rather than the discrete fixed blocks characteristic of current signal systems. Moving blocks thus have the potential to reduce minimum headways. With a fixed-block system, trains outside of terminals or interlocking limits traveling at the normal track speed are typically separated by at least two blocks, irrespective of their individual stopping characteristics. In contrast, in a moving block system, trains can be separated by little more than a single block and potentially by a distance related to the individual stopping distance of each train. This effectively reduces the minimum train separation from two or more blocks, as required with a fixed-block system, to a single block (or even less for some trains) of separation.

This is probably the greatest opportunity for enhanced capacity attributable to CBTC, especially on lines on which there is traffic with similar speeds but heterogeneous stopping distances. With a fixed-block system, the signals are spaced for the train with the longest braking distance, and therefore, the headway is longer than needed for much of the traffic. Slower or lighter trains with shorter braking distances, such as passenger and commuter trains, will be able to follow other train traffic more closely. This might help mitigate the disproportionate impact of certain types of heterogeneity due to the mixing of passenger and freight traffic (*3–6*).

Moving blocks also offer a benefit during recovery from temporary track outages or delays. Successive trains will be able to follow each other more closely because of their shorter braking distances at slower speeds. With a single track, to get operations back to normal as quickly as possible, moving blocks will allow trains to be fleeted through the work area with much closer spacing than is possible with conventional signal systems. This fleeting may also be of value when a double-track section must be made into a single track during maintenance (*24*).

Moving block capability can also reduce the delays due to passes on single-track lines. Shorter headways reduce the time that the overtaken train waits in the siding (*33*). Also, when a train leaves the siding, new movement authority can be issued to a train immediately after an overtaking train has passed the exit switch and the switch has been lined. It is not necessary to wait until the first block has been cleared, as may sometimes be required with conventional traffic control systems (*24*).

ELEMENTS OF AN ECP BRAKE SYSTEM THAT WILL AFFECT CAPACITY

ECP brakes change how the brake signal is transmitted. The signal will be transmitted through an electronic signal instead of a reduction in train line air pressure. Currently, each car is connected with an air line that is used to charge the brakes and transmit the braking signal. With ECP brakes, each car will also be connected by an electrical connection.

Current Systems

The current pneumatic brake system uses air pressure both to transmit the braking signal and to charge the brake reservoirs of the cars in the train. A reduction in air pressure along the brake line causes

the control valve to admit air into the brake cylinder applying the brakes. Two important limitations in this system in typical North American freight train applications are that the system does not permit the reservoirs to be recharged while the brakes are being applied, and it does not permit graduated release. Repeated application and release of the brakes can deplete the air pressure in the reservoirs and substantially reduce the braking force available. Avoiding this poses several operational limitations that affect capacity and potential safety problems if the brake system is not handled properly. The other limitation is that the air pressure signal is transmitted along the length of the train at approximately two-thirds the speed of sound (*8*). With longer trains there is a lag between the time of application and release at the rear of the train and that at the front of the train, causing significant in-train forces. Consequently, this means that there is a direct relationship between propagation time and braking distances. This problem is reduced when distributed power (DP) is used because it permits the braking signal to be initiated at more locations in the train, thereby reducing the brake signal propagation time and, thus, the braking distance (*35*). Railroads are increasingly using DP, and one major railroad estimates that 50% of its operation now uses DP.

Elements of an ECP Brake System

ECP brakes have several characteristics that have the potential to affect capacity. These are

- Instantaneous transmission of the brake signal,
- Steady brake line pressure, and
- Self-monitoring capabilities.

Use of an electronic signal instead of air pressure to transmit the brake signal allows virtually instantaneous transmission, enabling the nearly simultaneous application or release of the brakes along the entire length of the train. ECP brakes have a steady brake pipe pressure that allows continuous charging of the brake reservoirs and charging even while the brakes are being applied. The use of a train line cable also allows real-time, self-diagnostic "health check" functions to be incorporated into the brake system; the information from these functions informs the train crew when maintenance is needed (*8*). Each of these characteristics will be considered for their impact on capacity. Several proposed elements of an ECP brake system, including tricouplers and the ability to remotely uncouple cars, have the potential to affect capacity. These have not been included in any of the systems that have been developed, and therefore, they are not considered in this analysis.

Instantaneous Transmission of Brake Signal

With current brake systems, there is a delay during the propagation of the brake signal, whereas this delay is eliminated with ECP brakes. It is estimated that this will reduce the braking distance by about 40% to 60% compared with the conventional braking distance (*8*). Because the headway between trains is limited by the safe braking distance, if ECP brakes are installed on all trains, such a reduction will permit closer train spacing, if the traffic control system can accommodate it. The alternative to shorter headways is the ability to travel at higher speeds with the same signal spacing (*36*). Another benefit to having all the brakes on a train apply simultaneously is the reduction of in-train forces, permitting longer trains. Fewer, longer trains free up train slots, thereby allowing additional traffic. However, DP can provide some of the same benefits of reduced braking distances and longer train lengths but not the reduction in signal spacing that ECP brakes provide. Consequently, in some instances, railroads are already deriving some of the benefit that this aspect of ECP brakes offers.

Steady Brake Line Pressure

Steady brake line pressure allows the continuous charging of the brake reservoirs. This facilitates greater use of the braking system and reduces the time lost waiting to recharge the brake line and reservoir pressure after an application. With conventional freight train brakes, once the engineer has selected a brake level, the braking force cannot be reduced without completely releasing and reapplying the brakes. Trains must sometimes travel with more braking force applied than is necessary, resulting in slower operations (*8*). The continuous charging of brake reservoirs enables the graduated release of brakes, offering greater braking flexibility. This will potentially allow a train to conform more closely to appropriate track speed limits and increase average speeds. Another benefit is the shorter restarting time after stops. With the current brake technology, in areas of descending grades, the auxiliary reservoirs on each car of the train must be recharged before a train restarts from a stop (*8, 15*). This is not necessary with ECP brakes, reducing the dwell time on routes with large grades.

Self-Monitoring Capabilities

Use of an electrical signal to control the brakes has the added benefit of enabling the transmission of brake condition data to the locomotive. The engineer can monitor the brake condition and immediately be informed of any failure in any car on the train. In response to these capabilities, FRA issued a new regulation that requires brake inspections to be performed every 3,500 mi instead of every 1,000 mi, as is required with conventional brakes (*14*). This potentially allows an ECP-brake-equipped intermodal train originating from the ports of Los Angeles–Long Beach in California to travel all the way to Chicago, Illinois, without stopping for routine brake tests. Similarly, ECP-brake-equipped coal trains will be able to make quicker deliveries from western coal fields to power plants in the eastern and southern states (*37*). This not only decreases the cycle times but may also reduce the congestion at the terminals where these inspections currently take place. To achieve these results, the reconfiguration of terminal points and the resulting expenditures may be required.

IMPACT OF CBTC AND ECP BRAKES ON CAPACITY

The potential impact of CBTC and ECP brakes on capacity will depend on the type of implementation of each system, the traffic mix, the track configuration, and the topography of the route. For CBTC there are three different possible implementations: as a nonvital or a vital overlay to an existing control system or as a stand-alone system

(*25*). When CBTC is implemented as a nonvital overlay, the underlying control system provides movement authority, but CBTC provides an additional, automatic backup to prevent unsafe conditions. When it is implemented as a vital overlay, both the underlying system and CBTC verify and convey authority. In a stand-alone system, CBTC plays the sole role of verifying, conveying, and enforcing authority (*25*). Nonvital and vital overlay systems will still require the use of the current signal system, whereas a stand-alone system will permit moving blocks. Whether a route has single or multiple tracks will also affect the impacts of these systems. A single-track route is constrained because of the need for meets and passes, whereas with a multiple-track route, headway may be a more important constraint. The topography of the route also affects train handling and, consequently, capacity.

CBTC Nonvital Overlay System

A CBTC overlay provides enforcement according to the requirements of PTC, in addition to the current signal and traffic control systems. This type of implementation makes use of the current signal and traffic control system, and therefore, closer train spacing is not possible in wayside signal territory. However, in unsignaled ("dark") territory, an overlay system provides a more effective means of train separation. Much like a signal system, the installation of CBTC would efficiently allow the closer spacing of trains, thereby increasing capacity. Conversely, enforcement braking will result in trains slowing down sooner than they might otherwise, thereby reducing capacity. With or without a signal system, a CBTC overlay does not provide movement authority, and therefore, the current methods for this will remain in place, limiting some of the benefits of the in-cab display. In Europe the overlay version of ETCS has been found to reduce network capacity (*38*). In North America, the potential impact on capacity will be the greatest on signalized, single-track lines on which enforcement has a greater effect because of the more frequent stops from meets and passes.

CBTC Vital Overlay System

A CBTC vital overlay system will have capacity constraints similar to those of an overlay system because of the inability to take advantage of moving blocks. However, with a vital overlay system, the signal, traffic control, and CBTC systems are interconnected and authorities can be issued immediately via the in-cab display of the locomotive. Capacity under a vital overlay system will generally be the same or slightly higher than under that a nonvital system.

CBTC Stand-Alone System

A stand-alone CBTC system permits the use of real-time train and location data, in-cab displays, and moving blocks and the benefits that they provide. However, the potential capacity losses of braking enforcement still apply. The greatest potential benefit will be on multiple-track routes, on which reduced headways offer the greatest advantage. If moving blocks are used, this is likely to more than offset any potential capacity losses due to enforcement braking, which will have a resultant benefit to capacity.

Impact of ECP Brakes on Capacity

In an ECP brake system, the brake signal is transmitted instantaneously, the brake reservoirs are continuously charged, and the frequency of brake inspections is reduced. ECP brakes provide the greatest benefit relative to current systems for trains on severe grades (*8*). Grades can be bottlenecks on a railroad network, and ECP brakes provide improved train handling and reduced dwell time while the trains are traveling on these grades. Capacity can be improved on single-track lines because less time is lost during stops, and capacity can be improved on multiple-track lines because shorter headways are possible. Shorter cycles and increased terminal capacity can be achieved as well because of a reduction in the number of intermediate brake inspections.

Impact of Combination of CBTC and ECP Brakes

Use of the combination of CBTC and ECP brakes may allow better exploitation of the benefits that each system offers. It has been suggested that the data from ECP brakes will increase the accuracy of the braking algorithms, thereby reducing the impact of enforcement braking. Both of these systems increase the information available, and in combination, the additional train data from ECP brakes can be transmitted to the dispatcher or other relevant groups via the CBTC data network. Effective use of this information will permit a railroad to plan and manage its operations more efficiently. A stand-alone CBTC system will take the greatest advantage of ECP brakes because the use of moving blocks will permit railroads to reduce headways, which ECP brakes permit, without the need to modify signal spacing. Because it will take time for all trains to be equipped with ECP brakes, a stand-alone system will permit those trains equipped with ECP brakes to follow more closely behind the trains ahead, thereby providing incremental capacity benefits before the entire railcar fleet has been equipped with ECP brakes. A related benefit of CBTC with a moving block is that it will offer flexibility in train spacing if the train mix changes on a line or as further improvements in brake system effectiveness lead to shorter stopping distances and potentially closer train spacing.

DISCUSSION OF RESULTS

CBTC and ECP brakes make the train, signal, and traffic control systems more "intelligent" (*29*). This allows the railroad to better plan and control train movements, increasing railroad efficiency and capacity. However, braking enforcement will not increase capacity and may reduce it (*13, 33*). As the implementation of these technologies is considered, unanswered questions on their net effect on capacity remain.

Although railroads are planning to implement overlay CBTC systems and are testing ECP-brake-equipped unit trains, technical challenges remain. Conservative braking algorithms and excessive communications delays within CBTC may reduce capacity. Moving blocks also have not yet been proven to be technically feasible in the North American operating environment. CBTC may permit the removal of existing signal systems; however, to date there is no practical alternative to track circuits for the detection of broken rails. If track circuit systems cannot be eliminated, it may not be possible or economically justifiable to invest in a stand-alone CBTC system. Some authors have argued that even if it is possi-

ble, it may not be advisable to implement a completely stand-alone system (*39*).

Even when a reduction in headways is possible, this may not translate into additional network capacity because of other capacity bottlenecks. Headway is just one factor influencing capacity; other operational and infrastructure factors may continue to constrain a route. Sidings, interlockings, yards, and junctions are fixed points in the network; and reduced headways will not improve these capacity constraints. Additionally, terminals are considered major bottlenecks in many railroad networks (*40*). Consequently, although there may be reductions in the over-the-road time due to the use of CBTC and ECP brakes, increases in line capacity may not improve network capacity if the principal constraints are the terminals.

When the impacts of these new technologies are calculated, it is necessary to understand how their potential capacity benefits compare with what can be obtained from current systems. With ECP brakes, the comparative benefits of DP need to be considered. With CBTC the current train control technology on a line will affect the potential benefits of the system. In areas where there is no signal system or signals are widely spaced, CBTC will likely increase capacity. However, many of the areas that are currently facing the greatest capacity constraints are urban areas, where the signals are closely spaced. Lastly, the incremental benefit of CBTC is dependent on the implementation; in some cases, there may be no benefit without the use of a stand-alone system.

CONCLUSIONS AND FUTURE WORK

The implementation of CBTC and ECP brakes will have a direct effect on capacity. This paper has attempted to consider each critical characteristic of these technologies with respect to their capacity. All CBTC implementation types with enforcement braking have the potential for a loss of capacity, but as CBTC systems become more fully integrated, the potential for capacity enhancement improves. ECP brakes will provide benefits in most operational scenarios because of the shorter braking distances that they allow. Furthermore, CBTC may enable one of the principal benefits of ECP brakes—shorter stopping distances—to be more effectively and efficiently taken advantage of. These results will tend to be route and network specific, so individual railroads will need to conduct these analyses to understand the effects on their own systems.

Future work will use simulation software and mathematical models to quantify these impacts under a variety of scenarios of interest. Tests are planned with one and two main track lines, flat and mountainous terrain, and homogeneous and heterogeneous traffic. A better understanding of which combinations of conditions result in a gain or a loss of line capacity can be used to calculate how the railroad network as a whole will be affected by these technologies.

ACKNOWLEDGMENTS

The authors acknowledge the valuable insights and critical comments provided by Bill Moore Ede, Larry R. Milhon, Gordon Mott, Corey Pasta, Randolph R. Resor, Michael E. Smith, and David F. Thurston. This project was sponsored by a grant from the Association of American Railroads Technology Scanning Program. Additional support for the first author's graduate study has been provided by a CN Railroad Engineering Research Fellowship at the University of Illinois at Urbana—Champaign. This paper is adapted from an earlier version that was included in the *Proceedings of the 2009 AREMA Annual Conference* (American Railway Engineering and Maintenance-of-Way Association). Permission was obtained to use portions of that earlier paper.

REFERENCES

1. *Transportation—Invest in Our Future: America's Freight Challenge.* AASHTO, Washington, D.C., 2007.
2. *National Rail Freight Infrastructure Capacity and Investment Study.* Association of American Railroads, Washington, D.C., 2007.
3. Harrod, S. Capacity Factors of a Mixed Speed Railway Network. *Transportation Research, Part E,* Vol. 45, No. 5, 2009, pp. 830–841.
4. Dingler, M. H., Y.-C. Lai, and C. P. L. Barkan. Impact of Train Type Heterogeneity on Single-Track Railway Capacity. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2117,* Transportation Research Board of the National Academies, Washington, D.C., 2009, pp. 41–49.
5. Mostafa, K. K. *Actual Track Capacity of a Railroad Subdivision.* PhD thesis. Department of Civil and Environmental Engineering, University of Illinois at Urbana–Champaign, 1951.
6. American Railway Engineering Association. Notes on the Determination of the Traffic Capacity of Single and Multiple Track Railways. *Proc., 1921 American Railway Engineering Association Conference,* Chicago, Ill., 1921, pp. 744–759.
7. *Positive Train Control Systems Economic Analysis.* Docket FRA-2006-0132, Notice No. 1 RIN 2130-AC03. FRA, U.S. Department of Transportation, 2009.
8. *ECP Brake System for Freight Service.* Final Report. FRA, U.S. Department of Transportation, 2006.
9. Algoma Central Railway, British Columbia Railway, Burlington Northern Railroad, CN Rail, CP Rail, Norfolk Southern, and Seaboard System Railroad. *Advanced Train Control Systems Operating Requirements.* Prepared for Railway Association of Canada, Ottawa, Canada, and Association of American Railroads, Washington, D.C., April 1984.
10. Detmold, P. J. New Concepts in the Control of Train Movement. In *Transportation Research Record 1029,* TRB, National Research Council, Washington, D.C., 1985, pp. 43–47.
11. *Implementation of Positive Train Control Systems.* FRA, U.S. Department of Transportation, 1999.
12. *Electronically Controlled Pneumatic Brake Revenue Service Tests.* Publication DOT/FRA/ORD-99/07. FRA, U.S. Department of Transportation, 1999.
13. Moore Ede, W. J., A. Polivka, J. Brosseau, Y. Tse, and A. Reinschmidt. Improving Enforcement Algorithms for Communications-Based Train Control Using Adaptive Methods. *Proc., 9th International Heavy Haul Conference—Heavy Haul Innovation and Development,* Shanghai, China, 2009, pp. 722–727.
14. *Class IA Brake Tests—1,000-Mile Inspection.* 49 CFR Part 232.207, 2004.
15. Blank, R. W., B. McLaughlin, and S. K. Punwani. ECP Heavy Haul Coal Operations in North America with ECP Braking. *Proc., 9th International Heavy Haul Conference—Heavy Haul Innovation and Development,* Shanghai, China, 2009, pp. 558–564.
16. FRA, U.S. Department of Transportation. *Positive Train Control Systems.* 49 CFR Parts 229, 234, 235, and 236, Jan. 15, 2010.
17. Lee, J.-D., J.-H. Lee, C.-H. Cho, P.-G. Jeong, K.-H. Kim, and Y.-J. Kim. Analysis of Moving and Fixed Autoblock Systems for Korean High Speed Railway. In *Computer in Railways VII.* WIT Press, Southampton, MA, 2000, pp. 842–851.
18. Wendler, E. Influence of ETCS on the Capacity of Lines. In *Compendium on ERTMS,* Eurail Press, Hamburg, Germany, 2009, pp. 211–223.
19. Smith, M. E., R. R. Resor, and P. Patel. Train Dispatching Effectiveness with Respect to Communication-Based Train Control: Quantification of the Relationship. In *Transportation Research Record 1584,* TRB, National Research Council, Washington, D.C., 1997, pp. 22–30.
20. Smith, M. E., and R. R. Resor. The Use of Train Simulation as a Tool to Evaluate the Benefits of the Advanced Railroad Electronics System. *Journal of the Transportation Research Forum,* Vol. 29, No. 1, 1989, pp. 163–168.

21. Smith, M. E., P. K. Patel, R. R. Resor, and S. Kondapalli. Quantification of Expected Benefits: Meet/Pass Planning and Energy Management Subsystems of the Advanced Railroad Electronics System (ARES). *Journal of the Transportation Research Forum,* Vol. 30, No. 2, 1990, pp. 301–309.
22. Resor, R. R., M. E. Smith, and P. K. Patel. Positive Train Control (PTC): Calculating Benefits and Costs of a New Railroad Control Technology. *Journal of the Transportation Research Forum,* Vol. 44, No. 2, 2005, pp. 77–98.
23. Martland, C. D., and M. E. Smith. Estimating the Impact of Advanced Dispatching Systems on Terminal Performance. *Journal of the Transportation Research Forum,* Vol. 30, No. 2, 1990, pp. 286–300.
24. Moore Ede, W. J. How CBTC Can Increase Capacity. *Railway Age,* April 2001, pp. 49–50.
25. Drapa, J., B. Moore Ede, and A. Polivka. PTC Starts to Roll out Across the USA. *Railway Gazette International,* May 2007, pp. 285–289.
26. Kull, R. C. ECP Braking and PTC for Increasing Heavy Haul Railway Capacity. *Proc., 9th International Heavy Haul Conference—Heavy Haul and Innovation Development,* Shanghai, China, 2009, pp. 551–557.
27. Dick, C. T. Impact of Positive Train Control on Railway Capacity. Unpublished Report. Department of Civil and Environmental Engineering, University of Illinois at Urbana–Champaign, 2000.
28. Twombly, J. Migration from Conventional Signaling to Next-Generation Train Control. In *Transportation Research Record 1314,* TRB, National Research Council, Washington, D.C., 1991, pp. 122–125.
29. Ditmeyer, S. R. Network-Centric Railroading: Utilizing Intelligent Railroad Systems. *Proc., AREMA 2006 C&S Technical Conference,* Louisville, Ky., 2006.
30. Hoelscher, J., and L. Light. Full PTC Today with Off the Shelf Technology: Amtrak's ACSES Overlay on Expanded ATC. *Proc., 2001 AREMA Conference,* Chicago, Ill., 2001.
31. Thurston, D. F. Signaling and Capacity Through Computer Modeling. *Proc., International Conference on Computer Modeling for Rail Operations,* Delray Beach, Fla., 2004.
32. Anderson, D. Study of the Sensitivity of Predicted Stopping Distances to Changes to Input Parameters. Transportation Technology Center, Association of American Railroads, Pueblo, Colo., 1995.
33. *North American Joint Positive Train Control Project,* Docket No. DOT/FRA.ORD-09/04. FRA, U.S. Department of Transportation, April 2009.
34. *Benefits and Costs of Positive Train Control.* FRA, U.S. Department of Transportation, 2004.
35. Barrington, B., and D. Peltz. 10,000 ft Distributed Power Intermodal Trains. *Proc., 9th International Heavy Haul Conference—Heavy Haul Innovation and Development,* Shanghai, China, 2009, pp. 673–681.
36. Carlson, F. G. An Experimental Electric Assist Brake System. In *Technology Digest.* Report TD94-021. Association of American Railroads, Washington, D.C., March 1994.
37. *Electronically Controlled Pneumatic Brakes Fact Sheet.* FRA, U.S. Department of Transportation. http://www.fra.dot.gov/downloads/PubAffairs/ECP%20Brakes%20FINAL.pdf. Accessed June 15, 2009.
38. Strategic Rail Authority. *National ERTMS Programme Team 2004/05 Progress Report.* National ERTMS Programme, London, 2005.
39. Baugher, R. W. PTC: Overlay, or Stand-Alone? *Railway Age,* May 2004, p. 68.
40. Dirnberger, J. R., and C. P. L. Barkan. Lean Railroading for Improving Railroad Classification Terminal Performance: Bottleneck Management Methods. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1995,* Transportation Research Board of the National Academies, Washington, D.C., 2007, pp. 52–61.

The views expressed in this paper are those of the authors and do not necessarily reflect those of any of the sponsors or individuals mentioned above.

The Railroad Operating Technologies Committee peer-reviewed this paper.

Force Characteristics of Longitudinally Coupled Slab Track Turnout on Bridges Under Temperature Action

JuanJuan Ren, Rui Xiang, and Xueyi Liu

With a longitudinally coupled slab track turnout on LeiDa Bridge on the Wuhan–Guangzhou passenger dedicated line in China as an example, an integral finite element model of a turnout (crossover)–slab track–bridge–pier was developed. This model considered two No. 18 welded turnouts with movable frogs in the form of crossover, longitudinally coupled slab track, bridges, and piers as one system. The computational results included the temperature force and the displacement regularity of the turnouts, slab track, bridges, and piers. Their effect parameters, such as the extensional stiffness of the track slab, the friction coefficient of the sliding layer, and the setting of the anchor point, were also investigated. The key findings are that (*a*) the additional longitudinal forces and the displacement of rails induced by temperature changes increased with reduced slab extensional stiffness, whereas the corresponding forces of the turnout components decreased significantly; (*b*) the longitudinal distortions of the track structure were insensitive to the failure of the sliding layer but posed negative impacts on the forces and anchor points of the piers; and (*c*) it was necessary to arrange anchor points on large-span bridges to prevent rail displacement and reduce the negative effects of forces on the piers.

Experiences with many high-speed railway lines in China and other countries show that slab track systems have significant advantages (*1, 2*). The method for the design of the turnout in a slab track structure on a bridge and the technology for the construction of turnouts have not been thoroughly studied, and few findings from practical experience with the design and construction of turnouts throughout the world are available (*1*). The design of the slab track in the turnout zone on bridges encounters not only the technical difficulties as a result of the slab track on the bridge but also the complicated interaction between the turnout and its supporting base because of displacement limitations, which create greater complexity. Many projects with welded turnouts on bridges that use the slab track technology will be laid on Chinese passenger dedicated lines (PDLs), so more in-depth research on this particular topic is required. Because of the successful opening and operation of the Beijing–Tianjin intercity railway line in China, on which the innovative Bögl slab track system was used, the longitudinally coupled slab track will be extensively applied because of the success of this new concept. This report introduces the use of a longitudinal slab track system for transition zones on bridges by building a theoretical model and analyzing the feasibility of its application.

J. J. Ren and X. Liu, School of Civil Engineering, and R. Xiang, College of Traffic and Transportation, Southwest Jiaotong University, Chengdu 610031, China. Corresponding author: J. J. Ren, renjuanjuan1983@hotmail.com.

Transportation Research Record: Journal of the Transportation Research Board, No. 2159, Transportation Research Board of the National Academies, Washington, D.C., 2010, pp. 85–90.
DOI: 10.3141/2159-11

STRUCTURAL INTEGRAL MODEL

Bridge decks, track slabs (including prefabricated slabs and the concrete base), and rails (or turnouts) have displacements because of temperature changes, and these displacements bring about a longitudinal interaction among the three parts. In addition, stock rails, guide rails, prefabricated concrete slabs (or turnout slabs), concrete bases, bridge decks, and piers are coupled with restriction components, such as fastening systems, cement asphalt (CA) mortar, turnout position-limited devices, turnout spacer blocks, anchor points in the bridge superstructure, fixed bridge supports, and so forth. Consequently, four parts—turnouts, the slab track, the bridge deck, and piers—should be considered a coupled system for determination of the forces and displacements of each element.

The main characteristics of the system are as follows (*3*):

1. The concrete base of the slab track is completely coupled longitudinally and extends to the rub plate on the subgrade to connect the end thorn. The setting of the end thorn at the end of the 50-m-long rub plate is expected to balance the residual longitudinal force in the concrete base.
2. The concrete base is separated from the track superstructure on the bridge by a sliding layer that comprises one geotextile layer and a film. This minimizes the amount of friction that arises from relative movement between the bridge beam and the load-bearing slab and ensures that these two layers of the structure are able to expand and contract independently.
3. An anchor point in the bridge superstructure above every fixed support is expected to bear the loads on the bridge superstructure caused by braking trains.

This paper uses the finite element method for analysis and, as an example, two No. 18 welded turnouts with movable frogs in the form of crossovers laid on the LeiDa Bridge on the Wuhan–Guangzhou PDL in China. A model of the turnout, slab track, bridge deck, and pier is shown in Figure 1. The model of the longitudinal forces and distortions was used, according to the mechanical characteristics of the system. The 800-m-long model consists of 600-m-long bridges and two 50-m-long rub plates, as well as two 50-m-long embankment segments on both sides of the bridge. The design of the LeiDa

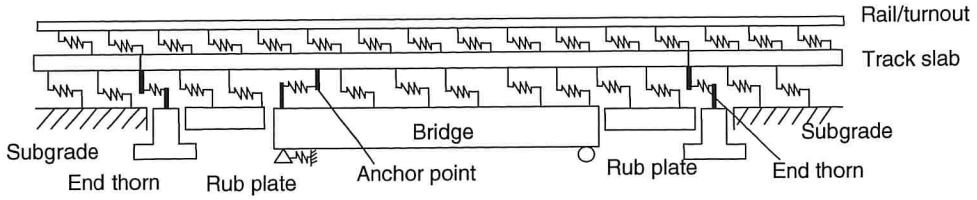


FIGURE 1 Model of system used for calculation.

Bridge spans consists of a one 24-m simple supported beam (SSB), two 32-m SSBs, a continuous beam (CB) of 40 + 64 + 40 m, and five 32-m SSBs plus a CB 6 × 32 m, as illustrated in Figure 2. On the basis of the design principle of welded turnouts on bridges (*4, 5*), this crossover is laid on a CB 6 × 32 m, and its center is above the fixed support in the middle of bridge. A schematic of the turnout–turnout plates–bridge and a cross section of the turnout zone on the bridge are illustrated in Figures 3 and 4, respectively. The following assumptions are applied in the model (*3*):

1. CA mortar works well and has normal bonding behavior so that the prefabricated slabs and the concrete base can deform in synchrony; the two layers can therefore be simplified into one layer, namely, a combined track slab (slab for short).
2. The rail, slab, and bridge are all simulated as beam elements.
3. Both the fastening system and the sliding layer inserted between the slab and the bridge are simulated as nonlinear springs.
4. Position-limited devices at the switch points and spacer blocks between short (or long) point rails and wing rails are simulated as nonlinear springs.
5. The anchor points, piers, and end thorns are all assumed to be linear springs; and the longitudinal stiffness of the anchor points and the end thorns are presumed to be infinite.

CALCULATION PARAMETERS

The model uses a Chinese-type (CHN) 60-kg/m rail and a Vossloh Loarv 300 (SKL 15, from Germany) rail fastener. The fastener has a slip resistance of 30 kN/m (unloaded) or 60 kN/m (loaded) after a displacement of 0.5 mm, which is determined according to DS 804 (German code). The 20-cm-thick, 2.55-m-wide prefabricated slabs are made of Grade C50 concrete, whereas the 18-cm-thick concrete base is made of Grade C40 concrete. The 24-cm-thick single-turnout slabs from Grade C50 concrete have various lengths, but the lengths do not exceed 6.5 m, and the widths are between 2.55 and 3.90 m. The track deadweight per unit length multiplied by the friction coefficient of the rub plate provides the slip resistance of the rub plate, and the track deadweight per unit length multiplied by the friction coefficient of the sliding layer provides the slip resistance of the sliding layer.

The longitudinal resistance values of position-limited devices and spacer blocks of No. 18 turnouts on the Qinhuangdao–Shenyang PDL were measured by China Railway Shanhaiguan Bridge Group Co., Ltd., Southwest Jiaotong University, and the China Academy of Railway Sciences in 2003 (*6*). The measured values were then used as parameters for the position-limited devices and the spacer blocks in the model used in the present study. The position-limited devices have a resistance of 1.5×10^5 kN/m within a relative displacement of

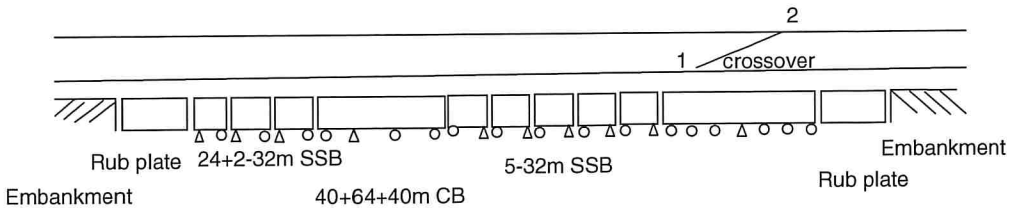


FIGURE 2 Arrangement of bridge spans and turnout location.

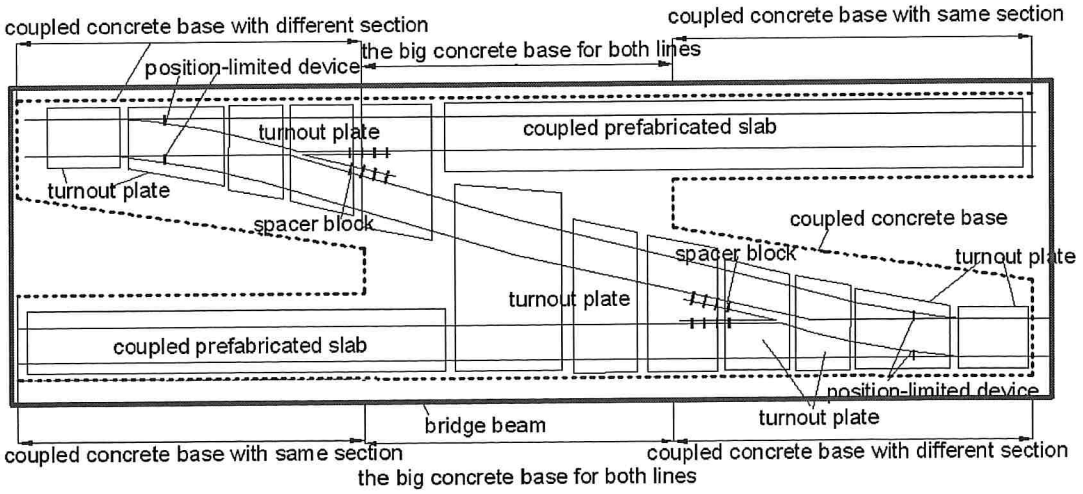


FIGURE 3 Schematic plan of longitudinally coupled slab track turnout on bridge.

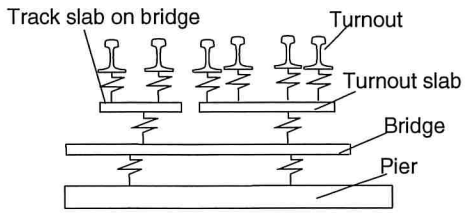


FIGURE 4 Cross section of turnout.

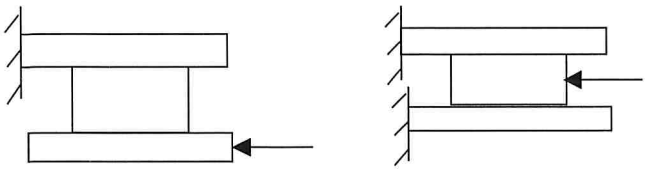


FIGURE 5 Resistance–displacement measurements of turnout spacer blocks.

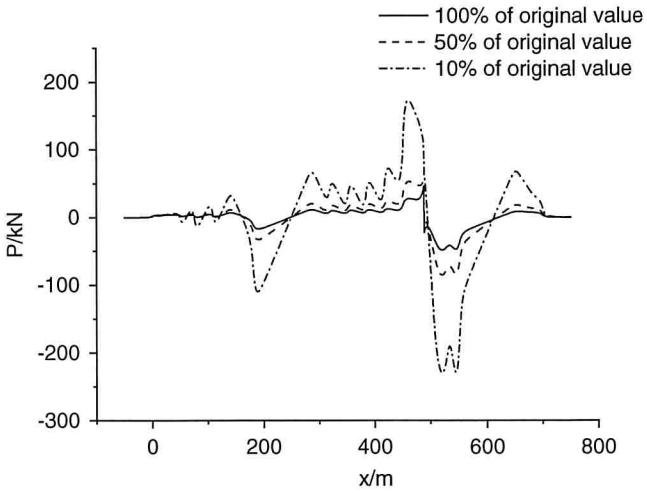


FIGURE 6 Forces on rail as result of temperature change (P = longitudinal temperature force).

1 mm and a resistance of 6×10^4 kN/m within a relative displacement of greater than 1 mm.

Two methods were applied to test the resistance of the spacer blocks. These are illustrated in Figures 5*a* and 5*b*. The results of the tests are shown in Table 1, which provides the resistance of the spacer block (which is equal to the load) and the relative displacements between the point rail and the wing rail.

By polynomial fitting, the resistance–displacement relationship of the spacer blocks can be written as the following function (F):

$$F(y) = 38.027 + 68.8y - 15.392y^2 + 2.333y^3 - 0.105y^4 \,(\text{kN}) \quad (1)$$

where $F(y)$ denotes the spacer block resistance and y denotes the relative displacement between the point rail and the wing rail.

EFFECTS OF TEMPERATURE CHANGES ON MODEL RESULTS

The temperatures of the rails, slabs, and bridge were reduced to 50°C, 30°C, and 20°C, respectively, to assess the impact of a temperature change on the model outputs (*7*).

Effect of Slab Extensional Stiffness

Temperature stress arises in the reinforced concrete slab because of temperature changes, resulting in cracks when the stress exceeds the ultimate tensile strength of the concrete. Hence, the extensional stiffness of the slab should be decreased to a certain extent by calculation. In the present study, slab extensional stiffness values of 100%, 50%, and 10% were applied in different cases; and the results are illustrated in Figures 6 and 7. Figures 6 and 7 show that forces on the rail as a result of temperature change and displacements decrease as the slab extensional stiffness increases. The incremental ratio of the force on the rail as a result of temperature change and displacement remains within 70% when the extensional stiffness decreases from 100% to 50% but exceeds almost 200% (the most unfavorable case) when the extensional stiffness decreases from 50% to 10%. The peak value of the longitudinal force for stock rail is located at the heel of the switch rail, because a position-limited device transfers a single force from the curved guide rail to the stock rail. The relative displacements between the rails and the slabs are less than 0.4 mm, much lower than the upper limit of 20 mm for turnouts on bridges required by BWG, a German company. The displacements of and the forces on both the rails and the slabs clearly increase in the area of the turnouts [with a distance of the calculation position to the left end of rub plate (x) of 468 to 632 m] compared with those in the common area.

Effect of Friction Coefficient of Sliding Layer

The sliding layer comprises one geotextile layer and a film to reduce the interaction between the bridge beam and the track slab and to ensure that the bridge's beams are able to expand and contract independently. Although partial longitudinal forces can be transferred by the anchor points from the track structure to the bridge, the spare forces will be absorbed by this sliding layer via its gliding state. It is concluded that the sliding layer plays an important role in force transfer, and thus, its failure should be taken into account in an actual simulation and in system design and computing. From the experience with the Beijing–Tianjin intercity railway line, a friction coefficient (μ) design value of 0.2 is used for the sliding layer, and a μ value of 1.0, which indicates failure, is the most unfavorable case. Results in accordance

TABLE 1 Relative Displacement Between Point Rail and Wing Rail (y) of Turnout Spacer Blocks (mm)

	Load (kN)							
	50	100	150	200	250	310	410	490
Eccentric load	0.26	1.05	2.79	4.12	5.53	7.00	9.00	10.55
Centric load	0.08	1.36	2.56	4.52	6.00	7.04	8.60	10.80
Average value	0.17	1.20	2.68	4.32	5.77	7.02	8.80	10.68

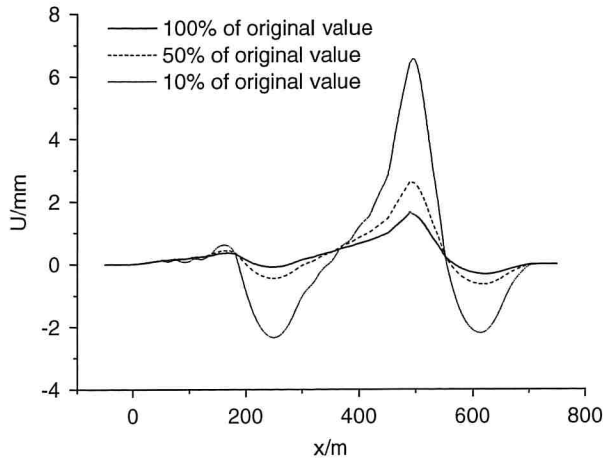


FIGURE 7 Rail displacements (U = rail longitudinal displacement).

with the two working states (a state in which the sliding layer works well and a state in which the sliding layer fails, both of which use a slab extensional stiffness of 100%), illustrated in Figures 8 and 9, show that the additional forces on the rail as a result of temperature change and rail displacements increase with a greater friction coefficient. This is because the interaction between the slabs and the bridge is enhanced by the greater friction coefficient for the sliding layer.

Effects of Anchor Points

Anchor points are longitudinally immovable connection points between the slab and the bridge and are created by setting dowels in the bridge decks to transfer the braking force into the bridge piers as soon as possible, and zero displacement between them is guaranteed. Calculation of the results for additional forces on the rail as a result of temperature change and displacements in different cases (with or without the setting of anchor points and with a slab extensional stiffness value of 100% or 10%) are presented in Figures 10 and 11. With the same slab stiffness and friction coefficient, it can be seen that additional forces on the rail and displacements increase if the anchor points are not arranged.

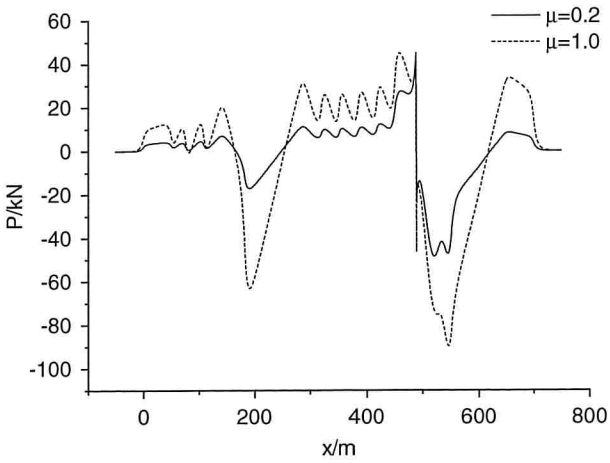


FIGURE 8 Additional forces on the rail as a result of temperature change.

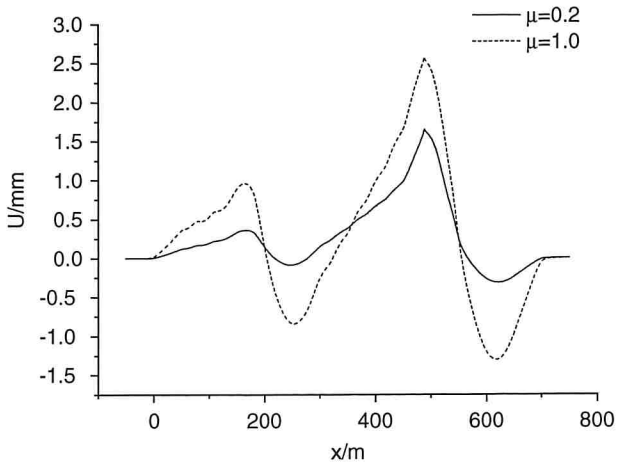


FIGURE 9 Stock rail displacements.

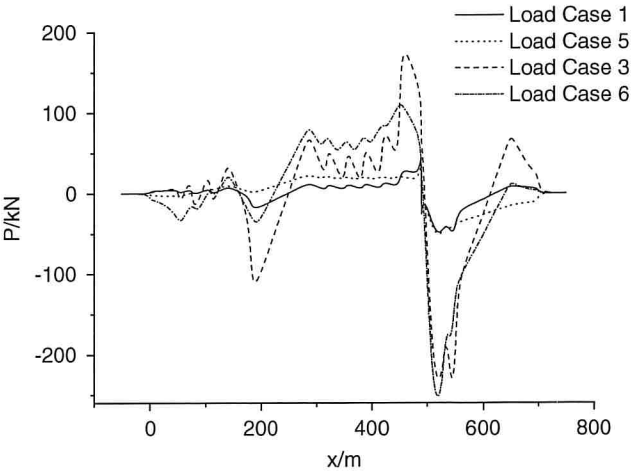


FIGURE 10 Additional forces on the rail as a result of temperature change.

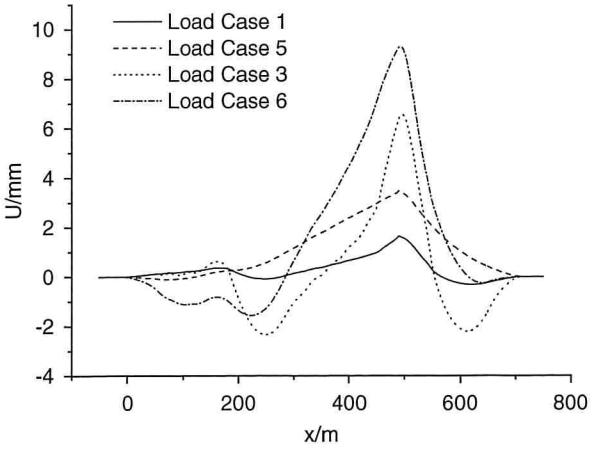


FIGURE 11 Rail displacements.

TABLE 2 Longitudinal Force of Piers in Different Load Cases (kN)

Load Case	Pier No. 0	1	2	3	4	5	6	7	8	9
1	−116.3	−40.1	−47.4	−224.7	−1.7	−20.5	−37.6	−78.5	−103.8	−80.4
2	−131.6	−45.7	−53.1	−256.6	11.7	−18.5	−48.6	−114.2	−164.0	−68.1
3	−152.4	−54.8	−57.4	−307.1	89.1	22.1	−47.9	−174.5	−321.9	28.0
4	−349.5	−115.9	−130.6	−572.5	30.3	−11.0	−53.4	−135.0	−192.9	−74.6
5	−230.9	−228.8	−222.5	−630.1	135.5	128.0	110.6	115.8	103.6	−77.7
6	−185.6	−166.9	−152.5	−506.8	133.7	83.0	32.8	−12.1	−58.7	−189.8

The longitudinal forces of the piers and turnout components as well as displacement of the switch point and the frog for the following six load cases are listed in Tables 2 to 4:

Load Case 1. With anchor points, slab stiffness of 100%, $\mu = 0.2$;
Load Case 2. With anchor points, slab stiffness of 50%, $\mu = 0.2$;
Load Case 3. With anchor points, slab stiffness of 10%, $\mu = 0.2$;
Load Case 4. With anchor points, slab stiffness of 100%, $\mu = 1.0$;
Load Case 5. Without anchor points, slab stiffness of 100%, $\mu = 0.2$; and
Load Case 6. Without anchor points, slab stiffness of 10%, $\mu = 0.2$.

From the results it can be concluded that

1. The longitudinal forces of piers increase significantly if the anchor points do not exist in the system (Table 2).
2. The absolute displacements of crossover switch points and frogs increase slightly (Table 3) if the anchor points do not exist in the system, whereas their relative displacements from the slab mostly remain unchanged.
3. The maximal longitudinal forces of force-transfer components of the turnouts (viz., position-limited devices and spacer blocks) have a strong relation to the extensional stiffness of the slab, and an obvious decrease in the forces for the turnout components occurs with a reduction of the extensional stiffness (compare the data for Load Cases 1 to 3 in Table 4).
4. Displacements of the switch points and frogs and the forces of the force-transferring components of the turnouts are insensitive to the friction coefficient of the sliding layer, whereas the forces of the bridge piers are sensitive to a certain extent (compare the data for Load Cases 1 and 4 in Tables 2 to 4); thus, to ensure that the bearing of bridges is secure, it is necessary to settle the anchor points on long-span bridges.

TABLE 3 Absolute Displacements of Crossover Switch Point and Frogs (in mm)

Load Case	Item: Switch Point AD 1#	Switch Point AD 2#	Frog AD 1#	Frog AD 2#
1	20.94	−21.23	9.03	−8.63
2	21.86	−22.26	9.30	−8.92
3	25.44	−26.69	10.31	−10.30
4	21.77	−22.09	9.30	−8.89
5	22.74	−21.62	10.72	−9.29
6	28.14	−26.18	12.56	−10.41

NOTE: AD = absolute displacement; 1# = typifies turnout No. 1 of crossover; 2# = typifies turnout No. 2 of crossover.

CONCLUSIONS

Because of the mechanical particularities of longitudinally coupled slab turnouts on bridges, the whole system should be considered a three-layer rail (turnout)–slab track–bridge model. Meanwhile, rails and track slab can work synchronously when fasteners with longitudinal clamping forces greater than the traditional forces are used. Therefore, the relative displacements between the rails and the track slabs stay in a small range (maximum of ≤0.2 mm), which meets the requirements of the limitations of the turnout position. The sliding layer, anchor points, and end thorns are critical to the performance of the whole track system. The longitudinally coupled slab turnout technique is thereby one potential solution for slab track in turnout zones on bridges for the construction of Chinese PDLs. From the analysis of the whole system, the following conclusions are drawn:

1. When cracks appear on track slabs, their extensional stiffness decreases accordingly. Thus, additional longitudinal forces on and the displacement of rails induced by temperature changes will increase with a reduction in the slab extensional stiffness, whereas the corresponding forces on the components of the turnout, such as position-limited devices and spacer blocks, will decrease significantly.

TABLE 4 Maximal Longitudinal Force of Force-Transferring Components of Turnouts

Load Case	Force (kN): Position-Limited Devices 1#	Position-Limited Devices 2#	Spacer Blocks 1#	Spacer Blocks 2#
1	73.20	70.54	209.18	208.06
2	65.13	60.21	198.86	196.76
3	35.88	18.10	160.50	152.93
4	59.65	52.75	198.02	196.17
5	69.98	77.60	208.07	211.22
6	11.15	26.67	163.29	164.54

2. The friction coefficient of the sliding layer has little impact on the longitudinal distortions of the track structure but greatly affects the forces on the piers.

3. The failure of anchor points will not result in visible changes to the additional rail force but will lead to greater rail displacement. Therefore, it is necessary to arrange the anchor points in long-span bridges so that they reduce the negative effects on the forces on the piers because of anchor point failure.

REFERENCES

1. Lichtberger, B. *Track Compendium.* Eurailpress Tetylaff-Hestra GmbH, Hamburg, Germany, 2005.
2. Tayabji, S. D., and D. Bilow. Concrete Slab Track State of the Practice. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1742,* TRB, National Research Council, Washington, D.C., 2001, pp. 87–96.
3. Ren, J. J., X. Y. Liu, H. S. Lin, and A. A. M. Shubber. Influencing Factors of Temperature Force and Displacement of Longitudinal Continuous Ballastless Seamless Turnout on Bridges for PDLs. *Proc., International Conference on Transportation Engineering 2009,* Chengdu, China, Vol. 1, 2009, pp. 207–212.
4. Wang, P., and X. Y. Liu. *Calculating Theory and Design Method of Welded Turnout.* Southwest Jiaotong University, Chengdu, China, 2007.
5. Bonaventura, C. S., A. M. Zarembski, J. W. Palese, and D. R. Holfeld. Increasing Speeds Through the Diverging Route of a Turnout Without Increasing Lead Length. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1863,* Transportation Research Board of the National Academies, Washington, D.C., 2004, pp. 61–67.
6. *Research Report of CWR and Welded Turnout Optimization.* Series Report of Railway Track Laboratory, School of Civil Engineering, Southwest Jiaotong University, Chengdu, People's Republic of China, March 2003.
7. Zeng, Z. P., and X. F. Chen. Analysis of the Influence of Some Parameters on the Temperature Force and Displacement of Welded Turnout on Simple-Supported Beam Bridge. *China Railway Science,* Vol. 27, No. 1, 2006, pp. 53–58.

The Railroad Track Structure System Design Committee peer-reviewed this paper.

Stress Dependence of Ultrasonic Guided Waves in Rails

Ivan Bartoli, Robert Phillips, Stefano Coccia, Ankit Srivastava, Francesco Lanza di Scalea, Mahmood Fateh, and Gary Carr

Most modern railways use continuous welded rails (CWRs). A major problem in these structures is the almost total absence of expansion joints, which can create severe issues such as buckling in hot weather and breakage or pulling apart in cold weather. To minimize these risks, CWRs are built by connecting track segments that are prestressed before welding. A related critical parameter is the rail neutral temperature (NT), which is defined as the temperature at which the net longitudinal force in the rail is zero. When the ambient temperature is higher or lower than the NT, the rail is under compression and tension, respectively. Knowledge of the NT provides a potential method for indirect measurement of the stress and load in the rail. Unfortunately, the measurement of the in situ stress (or NT) has been a long-standing challenge for railway owners and operators. This paper presents numerical results on the dynamic behavior of CWRs subjected to a static axial stress. The results show how ultrasonic guided waves are sensitive to variations in stress and could potentially be used to estimate the stress level or the NT in rails. The present work represents the initial concept phase of a research and development study funded by the FRA. The ultimate objective of this study is to develop and test a prototype system that uses noncontact dynamic sensing to measure in situ rail stress in motion at speeds up to 30 mph to determine rail NTs and the related incipient buckling risks in CWRs.

According to FRA safety statistics data, rail buckling was responsible for 48 derailments and nearly $30 million in cost during 2006 alone in the United States. Buckling occurs when the high ambient temperatures generate large compression stresses in continuously welded rails (CWRs) that ultimately make the rail unstable. Information on the in situ stress level in the rail could prevent buckling. Alternatively, knowledge of the neutral temperature (NT), defined as the temperature corresponding to a stress-free condition in the rail, would provide an estimate of the critical temperature corresponding to the critical load. Hence, NT provides an indirect measurement of the stress and load in the rail, as follows (*1*):

$$P = \alpha EA(T - \mathrm{NT}) \tag{1}$$

where

P = applied load,
α = coefficient of thermal expansion of steel,
E = Young's modulus of steel,
A = rail cross-sectional area, and
T = ambient temperature.

Unfortunately, the measurement of applied stress (or NT) represents an unsolved problem for railway owners and operators. Measurement methods under consideration today include

- Measurement of static rail stiffness. Although measurement of static rail stiffness is effective, it is cumbersome and not practical because it requires the unfastening of ~100 ft of rail.
- Measurement of dynamic resonance of torsional mode of vibration (D'stresen method). The D'stresen technique is based on the measurement of dynamic resonance (acceleration amplitude) below 90 Hz for the torsional mode of vibration of the rail (*2*). The approach does not require unfastening of the rail, hence its potential attractiveness. Unfortunately, the method is highly sensitive to rail fastening and support conditions. Consequently, normal tie-to-tie variations can make the stress or NT measurement unreliable by use of this method.
- Measurement of ultrasonic velocity (acoustoelastic method). The acoustoelastic method for stress measurement in rails has been known for more than 30 years (*3*). Acoustoelastic stress measurement is based on the theory of finite deformations that produce a change in ultrasonic velocity with applied stress. It typically uses longitudinal, shear, or surface (Rayleigh) waves in the approximately megahertz frequency range. The biggest challenge of the technique is that the acoustoelastic variation of wave velocity with stress is extremely small (~0.1% velocity change per gigapascal of stress). This low sensitivity often masks the stress indications provided by other parameters that affect wave velocity (namely, temperature variations and steel microstructure variations). For this reason, the acoustoelastic method for stress measurement in rails has historically been challenging to implement under practical field conditions.

The railroad industry would benefit from the availability of a noninvasive technique for the measurement of stress in rails in a noncontact manner while the measurement device is in motion. The technique would require sensitivities large enough to overcome the effects of tie-to-tie variations, changing temperatures, and changing steel microstructure and would benefit both the safety and the economy of operation of railroads.

Ultrasonic guided waves, used in the past for the detection of defects in rails and other structural components with waveguide

I. Bartoli, R. Phillips, S. Coccia, A. Srivastava, and F. Lanza di Scalea, Nondestructive Evaluation and Structural Health Monitoring Laboratory, University of California, San Diego, 9500 Gilman Drive, M.C. 0085, La Jolla, CA 92093-0085. M. Fateh and G. Carr, Office of Research and Development, FRA, 1200 New Jersey Avenue, SE, Washington, DC 20590. Corresponding author: I. Bartoli, ibartoli@ucsd.edu.

Transportation Research Record: Journal of the Transportation Research Board, No. 2159, Transportation Research Board of the National Academies, Washington, D.C., 2010, pp. 91–97.
DOI: 10.3141/2159-12

geometry (*4, 5*), have recently been considered to investigate their potential for stress measurement. At least two studies have predicted changes in the dispersion properties (velocity versus frequency, wave number versus frequency) of guided waves as a function of rail stress (*6, 7*). The present work studies such an effect numerically and analyzes changes in the time history signatures of transient ultrasonic guided waves in rails induced by preexisting axial stress.

Guided waves provide large monitoring ranges and complete coverage of the waveguide cross section. However, these waves can be fully exploited only once the complexities of wave propagation are managed. These complexities include the existence of multiple modes propagating simultaneously, frequency-dependent velocities (dispersion), and frequency-dependent attenuation. A semianalytical finite element method (FEM) was used in this study to predict the (guided wave) dynamic responses of CWRs under different axial stress levels.

SEMIANALYTICAL FINITE ELEMENT METHOD FOR MODELING OF GUIDED WAVE PROPAGATION IN RAILS

Semianalytical finite element (SAFE) methods have emerged as potentially viable tools for modeling the high-frequency wave propagation in waveguides of arbitrary geometries for which theoretical solutions are nonexistent or three-dimensional (3-D) finite element analysis (FEA) becomes computationally too intensive (*7–9*). The general SAFE approach for extracting the wave solutions (velocity, attenuation, mode shapes) uses a finite element (FE) discretization of the cross section of the waveguide alone. The displacements along the wave propagation direction are described theoretically as harmonic exponential functions. Thus, only a two-dimensional (2-D) FE cross-sectional discretization is needed and provides considerable computational savings compared with the computational ability required for a full 3-D FEA of the entire waveguide. In addition, because polynomial approximation of the displacement field along the waveguide is avoided, the method is applicable to the prediction of waves with very short wavelengths, for which a traditional 3-D approximation may fail.

The displacement field is assumed to be harmonic along the propagation direction, z, and shape functions, $N_k(x, y)$, are used to describe its amplitude in the cross-sectional plane of the rail ($x - y$) (Figure 1):

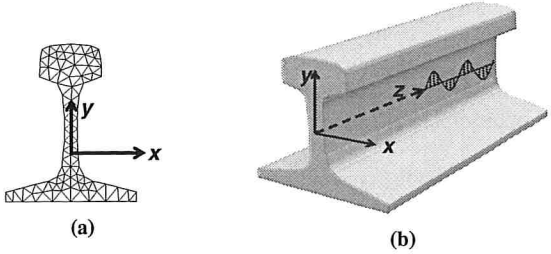


FIGURE 1 SAFE model for wave propagation in rails: (*a*) discretization of the rail cross section by 2-D FEs and (*b*) assumption of harmonic motion along propagation direction.

$$\mathbf{u}^{(e)}(x, y, z, t) = \begin{bmatrix} \sum_{k=1}^{n} N_k(x, y) U_{xk} \\ \sum_{k=1}^{n} N_k(x, y) U_{yk} \\ \sum_{k=1}^{n} N_k(x, y) U_{zk} \end{bmatrix}^{(e)} e^{i(\xi z - \omega t)} = \mathbf{N}(x, y)\mathbf{q}^{(e)} e^{i(\xi z - \omega t)} \quad (2)$$

where

$\mathbf{u}$ = element displacement vector;
(e) = generic element;
t = time variable;
n = total number of FEs;
k = wave number;
N = shape function matrix;
U_{xk}, U_{yk}, U_{zk} = nodal displacement components along the x, y, and z directions, respectively;
$\mathbf{q}$ = vector of the nodal displacement components; and
i = imaginary unit.

Harmonic motion [$e^{i(\xi z - \omega t)}$] is imposed along the longitudinal direction for a mode with frequency ω, which is equal to $2\pi f$, and wave number ξ, which is equal to $2\pi/\lambda$ (where λ is the wavelength).

The displacement assumption in Equation 2 can be introduced into the equations of motion. With standard mathematical manipulations, the following system of equations is finally obtained (*8, 9*):

$$\left[\mathbf{K}_1 + k\hat{\mathbf{K}}_2 + k^2\mathbf{K}_3 - \omega^2\mathbf{M}\right]\hat{\mathbf{U}} = \mathbf{f} \quad (3)$$

where mass matrix $\mathbf{M}$ is derived from the kinetic energy, and stiffness matrices $\mathbf{K}_i$ ($i = 1, 2, 3$) are derived from the strain energy. The vector $\hat{\mathbf{U}}$ represents the displacements, and $\mathbf{f}$ is the vector of nodal forces applied to the waveguide. The equation system can be rearranged as

$$(\mathbf{A} - \xi\mathbf{B})\mathbf{Q} = \mathbf{p} \quad (4)$$

where

$\mathbf{A}$ and $\mathbf{B}$ = matrices that depend on the stiffness and mass of the waveguide,
$\mathbf{Q}$ = displacement vector, and
$\mathbf{p}$ = loading terms.

The system of equations can be solved at separate frequencies because the terms are frequency dependent.

Unforced (Modal) Solutions

Additional information can be extracted from Equation 4. For example, by solving the eigenvalue problem under the assumption of a zero load ($\mathbf{p} = 0$), the dispersion curves of the guided modes propagating in a certain frequency range can be extracted. At each frequency f^*, a different eigenvalue problem,

$$(\mathbf{A} - \xi\mathbf{B})\mathbf{Q} = \mathbf{0} \quad (5)$$

where f is equal to f^*, is solved to obtain the wave numbers ξ_m (where m represents the mth mode) of all the propagating guided modes at f^*. The computation of the phase velocities (c_{ph}) can then

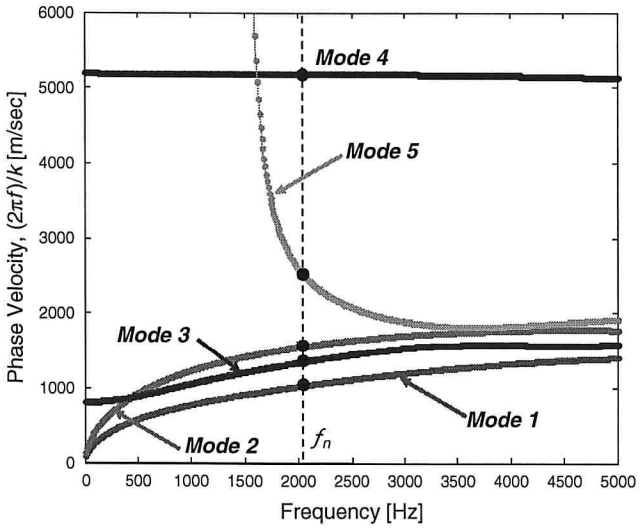


FIGURE 2 Phase velocity dispersion curves for a 115-lb AREMA rail section. The following properties were assumed: density ρ = 7,850 kg/m^3; Young's modulus E = 209 GPa; Poisson's ratio ν = 0.3.

be obtained as $c_{ph} = (2\pi f)/\xi_m$. Figure 2 shows the phase velocity dispersion curves for a 115-lb American Railway Engineering and Maintenance-of-Way Association (AREMA) rail section. It can be noted that by solving the eigenvalue problem in Equation 5 at a frequency $f = f^* = f_n$ = 2,000 Hz, as many as five propagating modes exist in the rail. Mode 1 is a flexural vertical mode, Mode 2 is a flexural horizontal mode, Mode 3 is a torsional mode, and Mode 4 is a longitudinal extensional mode. The only nonfundamental mode (a mode that does not propagate at zero frequency) is Mode 5, which has a cutoff frequency above 1,500 Hz. The mode shapes of the four fundamental modes will be shown in the following sections.

Forced Solutions

Once Equation 5 is solved for each frequency in the range of interest, the SAFE method can be used to quickly compute the rail response due to a general transient excitation. The forced solution for a load applied at z equal to z_S (**U**) is computed at each frequency as a combination of all the propagating modes by using the following expression:

$$\mathbf{U}(x,f) = \sum_{m=1}^{2M} -\frac{\Phi_m^L \tilde{\mathbf{p}}}{B_m} \Phi_m^{Rup} e^{i[\xi_m(z-z_S)]} \tag{6}$$

where

M = size of the system of equations,
m = mth mode,
up = first M elements of the vector Φ_m^R that has $2M$ elements and a total length of $2M$,
Φ_m^L = left eigenvector computed from the eigenvalue problem in Equation 5,
Φ_m^R = first half of the right eigenvector,
$\tilde{\mathbf{p}}$ = amplitudes of the nodal loads, and
matrix $B_m = \Phi_m^L \mathbf{B} \Phi_m^R$.

Equation 6 represents the response to a pure harmonic excitation of unitary amplitude. The response to an excitation with an arbitrary time history (**V**) can then be computed in the following manner. First, the frequency content of the excitation signal, $F(t)$, is computed by applying the Fourier transform ($\hat{F}$):

$$\hat{F}(f) = \int_{-\infty}^{\infty} F(t) e^{-i2\pi ft} dt \tag{7}$$

The response in the frequency domain to force $F(t)$ can then be computed by convolution as

$$\begin{aligned} \mathbf{V}(x,f) &= \hat{F}(f) \cdot \mathbf{U}(x,f) \\ &= \hat{F}(f) \cdot \sum_{m=1}^{M} -\frac{\Phi_m^L \tilde{\mathbf{p}}}{B_m} \Phi_m^{Rup} e^{i[\xi_m(z-z_S)]} \end{aligned} \tag{8}$$

Finally, the time-domain response can be obtained by using the inverse Fourier transform:

$$\mathbf{V}(x,t) = \int_{-\infty}^{\infty} \mathbf{V}(x,f) e^{i2\pi ft} df \tag{9}$$

Equation 9 is used in the next section to predict the forced response of a rail because of a narrowband tone burst excitation.

CHANGES IN WAVE VELOCITY IN RAILS UNDER AN APPLIED STRESS LEVEL

The application of a load to an elastic waveguide as the CWR can influence the propagation of the wave in the waveguide.

Regular FEM techniques can be used to analyze the change in wave propagation properties with the applied axial load in a waveguide (*6*). The SAFE method represents a more general and computationally efficient approach; the method was recently extended to include an initial axial load (*10*), and the dispersion curves were obtained for infinitely long waveguides. It was also shown that the addition of a single term to the expression of the linear strain energy is sufficient to account for the axial load applied on the waveguide along the longitudinal, z, direction. In such case, Equation 3 becomes

$$\left[\mathbf{K}_1 + \xi \hat{\mathbf{K}}_2 + \xi^2 \left(\mathbf{K}_3 + \mathbf{K}^{(0)}\right) - \omega^2 \mathbf{M}\right]_M \hat{\mathbf{U}} = \mathbf{f} \tag{10}$$

where the new additional stiffness matrix, $\mathbf{K}^{(0)}$, is proportional to mass matrix **M** and to the applied axial stress, $\sigma_{zz}^{(0)}$:

$$\mathbf{K}^{(0)} = \frac{\sigma_{zz}^{(0)}}{\rho} \mathbf{M} \tag{11}$$

where ρ is density.

In the present study, the SAFE method was used to identify changes in wave velocity due to different stress levels in 115-lb AREMA rails. Figure 3 shows the phase velocity dispersion curves of the four dominant vibrational wave modes for the case of an unloaded rail (Figure 3*a*) and a rail loaded axially at 0.1% tensile strain (Figure 3*b*) obtained by the SAFE method. Appreciable changes in wave velocity between the unloaded and the loaded rails can be observed for the flexural vertical, the flexural horizontal, and the torsional wave modes in the midfrequency (approximately kilohertz) range. It should be noted that this change in velocity is much higher than the

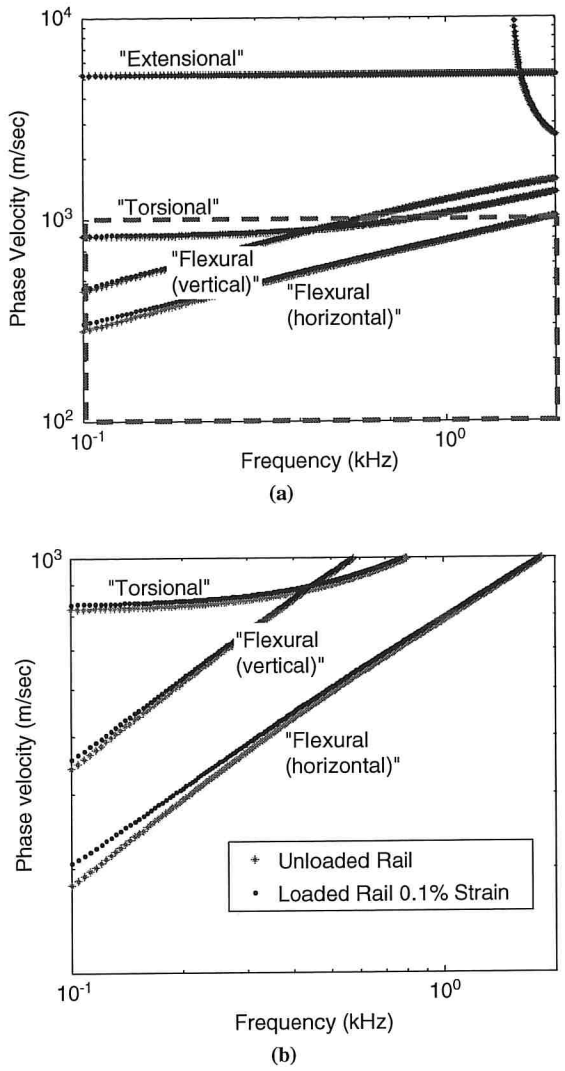


FIGURE 3 Phase velocity dispersion curves of the four dominant wave modes for case of (*a*) unloaded rail and (*b*) rail loaded axially at 0.1% tensile strain.

one obtained by the conventional acoustoelastic method, which uses frequencies in the megahertz range.

Figure 4 shows the mode shapes of the two fundamental flexural modes. As expected, the flexural horizontal mode has large lateral (x-axis) displacement components (Figure 4*a*), whereas the flexural vertical mode has a prevalent vertical (y-axis) motion (Figure 4*b*). Figure 5 shows the torsional mode, which is also influenced by the applied axial load. Figure 6 shows the longitudinal (extensional) mode, which is, instead, unaffected by the axial load.

The information provided by this analysis could be useful for the development of a prototype able to detect stress levels in the rail. On the basis of these results, the wave generation and detection schemes of the prototype design can be tailored to maximize the ultrasonic energy of the stress-sensitive modes while minimizing the energy of the stress-insensitive ones. It can be noted that all the modes considered have a nonnegligible displacement in the base of the rail. Such displacement is generally influenced by constraints (ties, fasteners, ballast), which can therefore affect the propagation velocities of these rail modes. However, some of the modes generate a smaller base motion than others. SAFE analysis can identify such modes that minimize base motion, hence increasing the robustness of the stress (or NT) measurement technique against tie-to-tie variations and other rail support variabilities.

The SAFE method was also used to study the dynamic response of the rail under different excitations. Equations 6 to 9 were used to compute the displacement in a generic location of the rail due to a point dynamic force. Figure 7 shows the numerical results obtained for the case of a transverse excitation on the side of the rail head. The seven-cycle tone burst (Figure 7*a*) is the force time history that induces the displacement time history shown in Figure 7*b*. The horizontal displacement is the result of two modes (flexural horizontal and torsional), both of which propagate at similar velocities. The results are shown for an unloaded rail and for a rail loaded at 0.1% of axial strain in tension. It can be seen that the velocity of the waves is indeed stress dependent, because the arrival times of the wave packets vary between the loaded and the unloaded rail conditions. This confirms the results shown by the dispersion curves in Figure 3. It should also be noted that the variation in arrival times in Figure 7*b* is large enough to be measured experimentally. Again, this is in contrast to the results of the acoustoelastic technique at approximately megahertz frequencies, which attempts to track much smaller velocity changes.

The case of a dynamic axial excitation of the rail is shown in Figure 8. It can be noted that this excitation generates two guided modes: the faster longitudinal mode and the slower flexural vertical mode. As discussed earlier in this paper, the longitudinal guided mode is unaffected by the stress variation, whereas the flexural mode becomes faster when the rail is under axial tension. This behavior is confirmed by the displacement time history results shown in Figure 8 (no change as a function of stress for the longitudinal wave arrival; appreciable changes for the vertical flexural wave arrival).

This type of analysis can be extended to track different stress levels in the rail, both in tension and in compression, in an effort to ultimately provide the information needed for the development of a rail stress (or NT) measurement prototype (wave modes, frequencies, excitation, and detection schemes).

CONCLUSIONS AND FUTURE RESEARCH

This paper describes part of the numerical results obtained in an investigation of the dynamic behavior of CWRs under changing axial stresses. The preliminary results show the changes in the propagation velocities of the axial, flexural vertical, flexural horizontal, and torsional wave modes of rails with an applied stress level (comparable to the yield stress of steel, or 1,000 microstrains). The features considered are in the kilohertz frequency range.

Currently planned research that will be reported in the future will focus on higher frequencies of guided waves, from which the increasing stress and incipient buckling cause nonlinear vibrations.

On the basis of these promising numerical results, University of California, San Diego, (UCSD), researchers plan to perform large-scale experimental tests at the UCSD Powell Structural Laboratories, where a test site with a 70-ft rail will be installed. The purpose of these tests will be to validate the dynamic parameters identified in the numerical models with respect to rail stress levels (hence, NT) and the related prediction of incipient buckling.

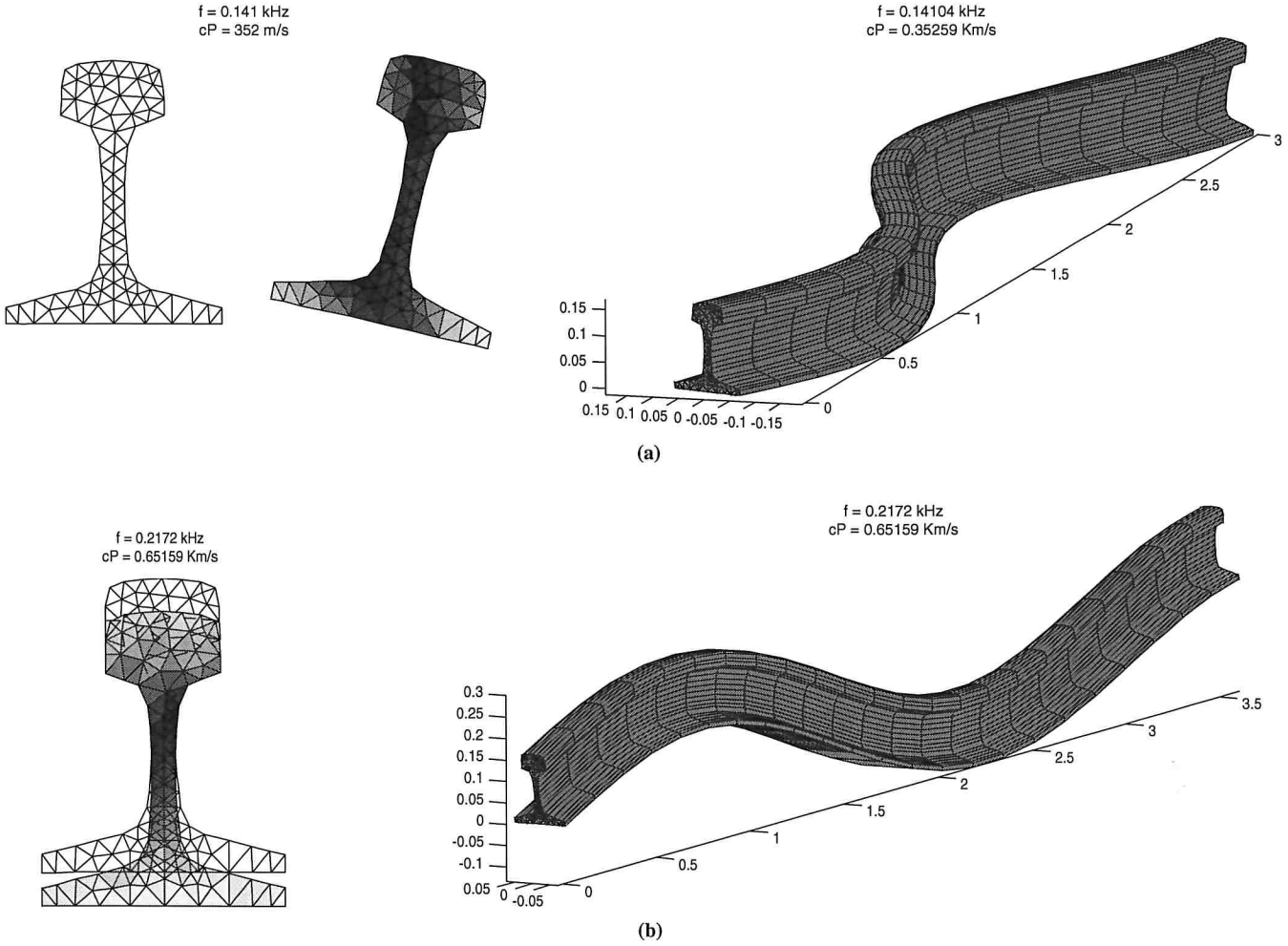


FIGURE 4 Guided modes sensitive to stress (mode shapes from SAFE analysis): (*a*) flexural horizontal mode and (*b*) flexural vertical mode (Cp = phase velocity).

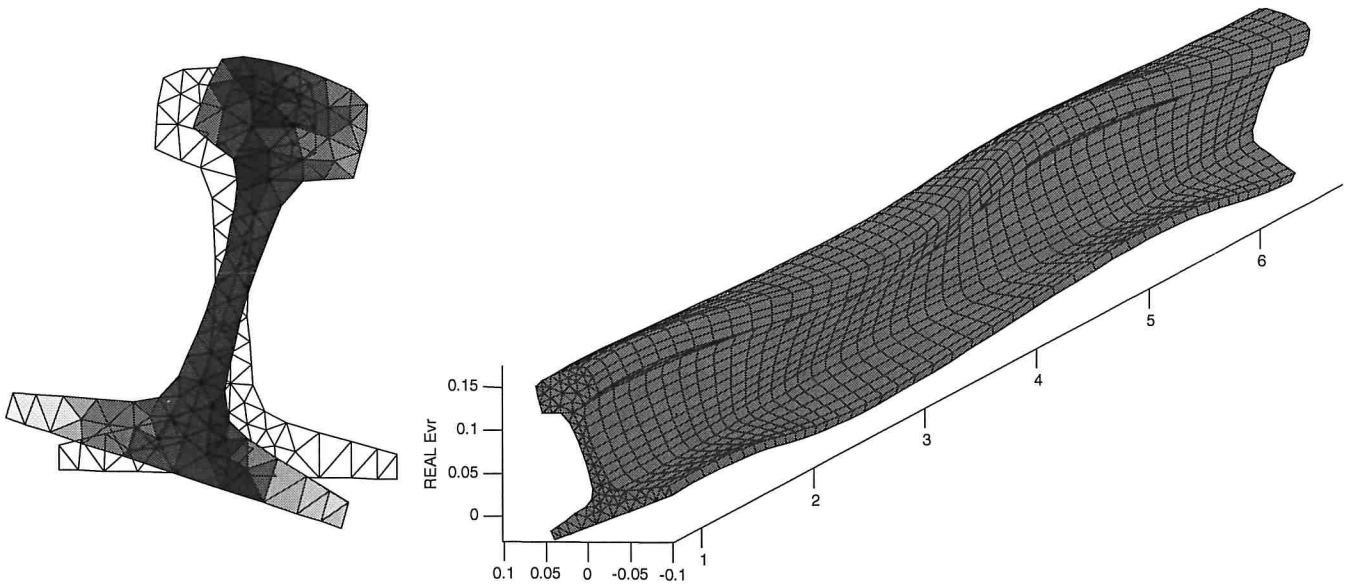


FIGURE 5 Guided modes sensitive to axial stress (mode shapes from SAFE analysis): torsional mode (Evr = Φ_m^R).

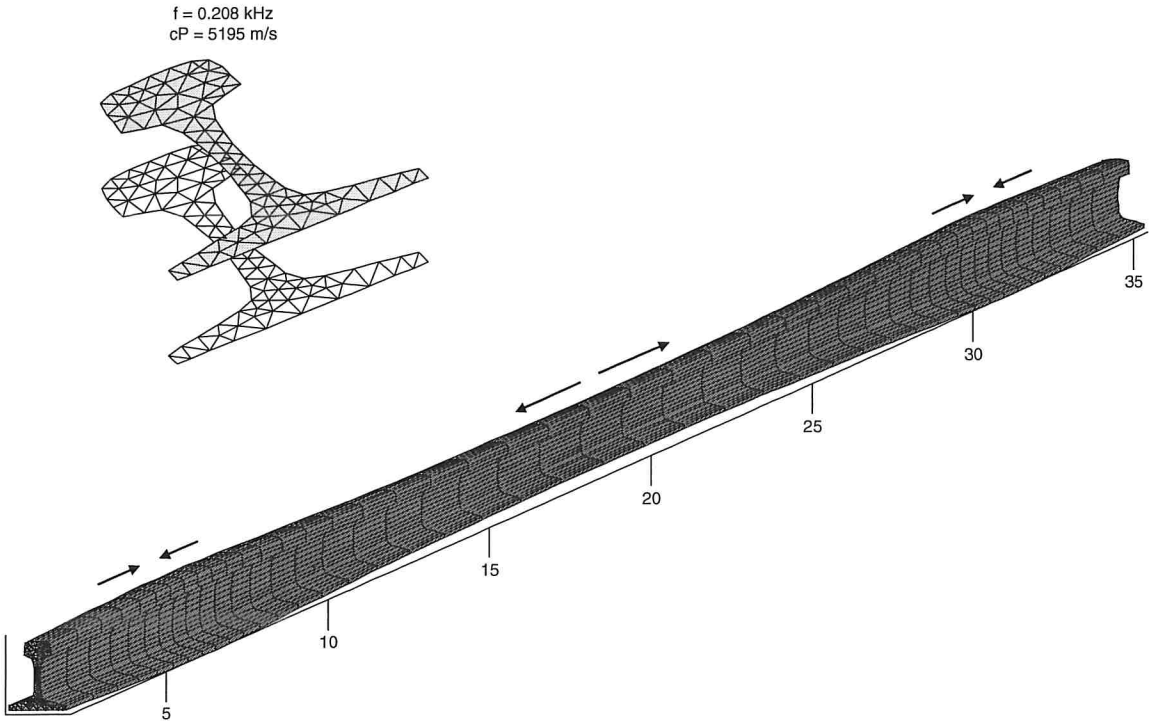
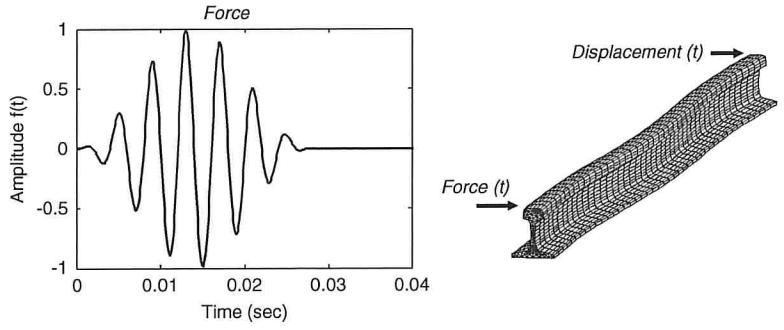
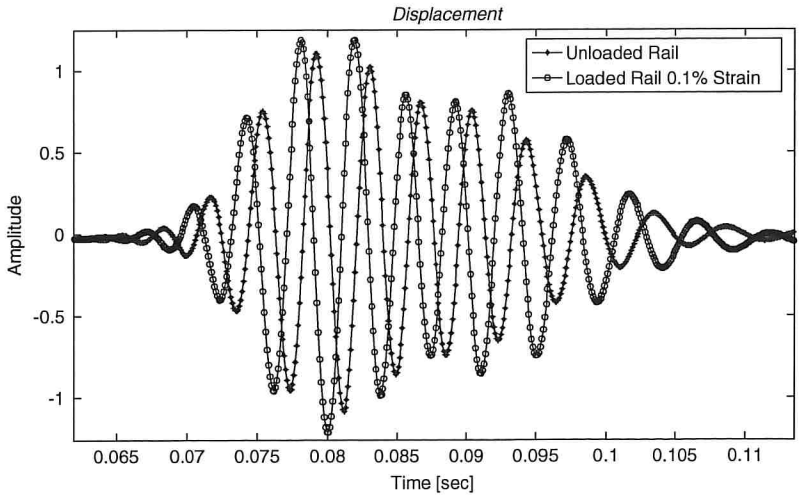


FIGURE 6 Guided modes insensitive to axial stress (mode shapes from SAFE analysis): longitudinal (extensional) mode.



(a)



(b)

FIGURE 7 Numerical results (time histories) obtained by SAFE method showing wave velocity change with applied stress: (*a*) narrowband excitation and (*b*) displacement time history for loaded rail and unloaded rail.

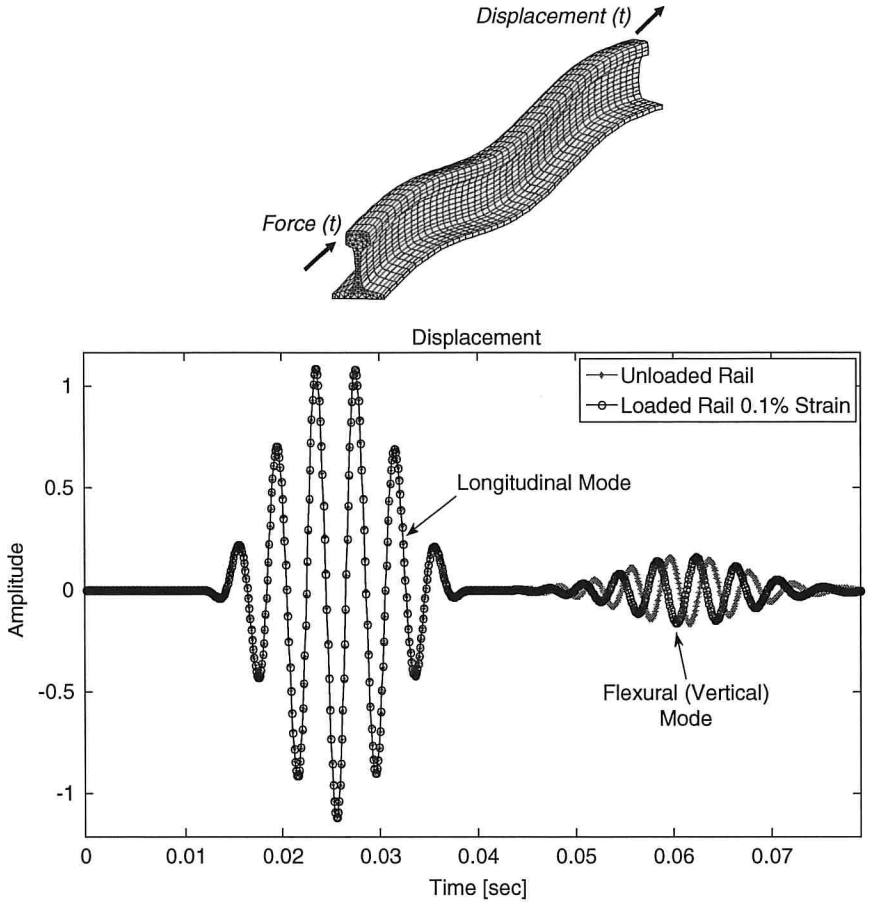


FIGURE 8 SAFE numerical results (time histories) of SAFE method showing wave velocity changes with stress: longitudinal (z-direction) displacement time history for loaded rail and unloaded rail.

ACKNOWLEDGMENT

This work was supported by FRA.

REFERENCES

1. Kish, A., and D. Clark. Better Management of CWR Neutral Temperature Through More Efficient Distressing. *Proc., 2004 AREMA Conference,* Nashville, Tenn., 2004.
2. Read, D., and B. Shust. Investigation of Prototype Rail Neutral Temperature Measurement System. *Railway Track and Structures,* June 2007, pp. 19–21.
3. Egle, D. M., and D. E. Bray. Measurement of Acoustoelastic and Third-Order Elastic Constants for Rail Steel. *Journal of the Acoustical Society of America,* Vol. 60, 1976, pp. 741–744.
4. Wilcox, P., M. Evans, B. Pavlakovic, D. N. Alleyne, K. Vine, P. Cawley, and M. J. S. Lowe. Guided Wave Testing of Rail. *Insight,* Vol. 45, 2003, pp. 413–420.
5. Rose, J. L., M. J. Avioli, P. Mudge, and R. Sanderson. Guided Wave Inspection Potential of Defects in Rail. *NDT&E International,* Vol. 37, 2004, pp. 153–161.
6. Chen, F., and P. D. Wilcox. The Effect of Load on Guided Wave Propagation. *Ultrasonics,* Vol. 47, 2007, pp. 111–122.
7. Damljanovic, V., and R. L. Weaver. Forced Response of a Cylindrical Waveguide with Simulation of the Wavenumber Extraction Problem. *Journal of the Acoustical Society of America,* Vol. 115, 2004, pp. 1582–1591.
8. Hayashi, T., W. J. Song, and J. L. Rose. Guided Wave Dispersion Curves for a Bar with an Arbitrary Cross-Section, a Rod and Rail Example. *Ultrasonics,* Vol. 41, 2003, pp. 175–183.
9. Bartoli, I., A. Marzani, F. Lanza di Scalea, and E. Viola. Modeling Wave Propagation in Damped Waveguides of Arbitrary Cross-Section. *Journal of Sound and Vibration,* Vol. 295, 2006, pp. 685–707.
10. Loveday, P. Semi-Analytical Finite Element Analysis of Elastic Waveguides Subjected to Axial Loads. *Ultrasonics,* Vol. 49, 2009, pp. 298–300.

The Railway Maintenance Committee peer-reviewed this paper.

Nondestructive Identification of Freezing-Induced Cracks in Concrete Sleepers of High-Speed Railways in South Korea

Sung-Ho Joh, Seon Keun Hwang, Tae-Ho Kang, Chul-Soo Park, and Il-Wha Lee

The construction of a high-speed railway in South Korea was recently impeded by the unexpected problem of cracks in sleepers. These were presumably induced by the volume expansion of freezing water entrapped in the fastening assemblies during the winter. A total of 306,000 sleepers had to be investigated on an urgent basis for hairline cracks and possible latent cracks. These findings suggested that an appropriate nondestructive testing (NDT) technique for the inspection of concrete sleepers should be designed to meet two requirements: speed and reliability in testing. Research was performed to identify an optimal NDT technique by using sleepers with three different crack conditions: without any cracks, with a surface-opening crack, and with a hairline crack and possible latent cracks. Flexural rigidity rather than material stiffness or depth to multiple reflections was chosen as a key parameter for assessment of the structural integrity of a sleeper. The results allowed the design of a two-stage testing procedure: the flexural rigidity assessment of concrete tracks by antisymmetric Lamb (FRACTAL) waves technique for preliminary scanning and spectral analysis of surface waves (SASW) tests for in-depth investigation. A total of 18,177 sleepers that required testing were examined using FRACTAL tests, and 387 of these sleepers were double-checked using SASW tests. Before the overall examination of sleepers, a series of statistical analyses were performed to establish the criteria for defining the quality of a sleeper. The results of the FRACTAL test agreed with those of the SASW test in identifying sleepers with hairline cracks or possible latent cracks.

A sleeper is an essential structural element for transfer of the load from a train to the subgrade. The structural integrity of a sleeper is an important safety parameter in train operations, especially high-speed trains, such as the South Korea Train eXpress (KTX). The type of sleeper adopted for KTX railways is concrete and is built in a concrete slab track system with various thicknesses, ranging from 0.11 to 0.13 m. Recently, some newly constructed sleepers have exhibited both surface-opening and latent cracks, as shown in Figure 1. A total of 332 sleepers with surface-opening cracks were discovered among the 306,000 sleepers installed for the second-stage construction of the KTX railway from Taegu to Gyeongju, South Korea. A team of investigators organized by the South Korean government concluded that conical cracks had developed in the sleepers because of volume expansion of the freezing water entrapped in the embedded fastening assemblies during the winter. Among the 306,000 sleepers, 18,177 were identified to be suspicious or damaged by freezing. The problem was then to identify sleepers with internal damage for repair. Sleepers with surface-opening cracks are easily detected by visual inspection. However, sleepers with hairline cracks or latent cracks are hard to detect. To make matters worse, the KTX railway construction schedule required the inspections to be finished in a couple of months.

The goal of the research described here was to design an optimal nondestructive testing (NDT) method for the identification of sleepers with hairline cracks or latent cracks. The impact–echo method (*1, 2*), the spectral analysis of surface waves (SASW) method (*3*), and the flexural rigidity assessment of concrete tracks by antisymmetric Lamb waves (FRACTAL) method (*4*) were used to investigate three typical sleepers: an undamaged sleeper, a sleeper with surface-opening cracks, and a sleeper with hairline cracks and possible latent cracks. The proposed optimal procedure involves two stages of testing: FRACTAL tests for preliminary scanning and SASW tests for in-depth investigation. The procedure was applied to all 18,177 of the suspicious sleepers, and some typical results are presented in this paper.

NONDESTRUCTIVE TECHNIQUES FOR ASSESSMENT OF STRUCTURAL INTEGRITY OF CONCRETE SLEEPERS

Impact–Echo Test

In the impact–echo test, stress waves from an impulse source are applied to the top surface of a platelike structure and are then reflected back from the bottom of the surface or from internal anomalies within the structure. The multiple reflections between the two plate boundaries are analyzed to determine the thickness of the structure or the

S.-H. Joh, Department of Civil Engineering, Chung-Ang University, Nae-Ri, DaeDeok-Myeon, AnSeong-Si, GyeongGi-Do 456-756, South Korea. S. K. Hwang and C.-S. Park, Vehicle and Track Research Team, and T.-H. Kang and I.-W. Lee, Track and Geotechnology Research Team, Korea Railroad Research Institute, Uiwang-City, GyeongGi-Do, South Korea. Corresponding author: S.-H. Joh, shjoh@cau.ac.kr.

Transportation Research Record: Journal of the Transportation Research Board, No. 2159, Transportation Research Board of the National Academies, Washington, D.C., 2010, pp. 98–109.
DOI: 10.3141/2159-13

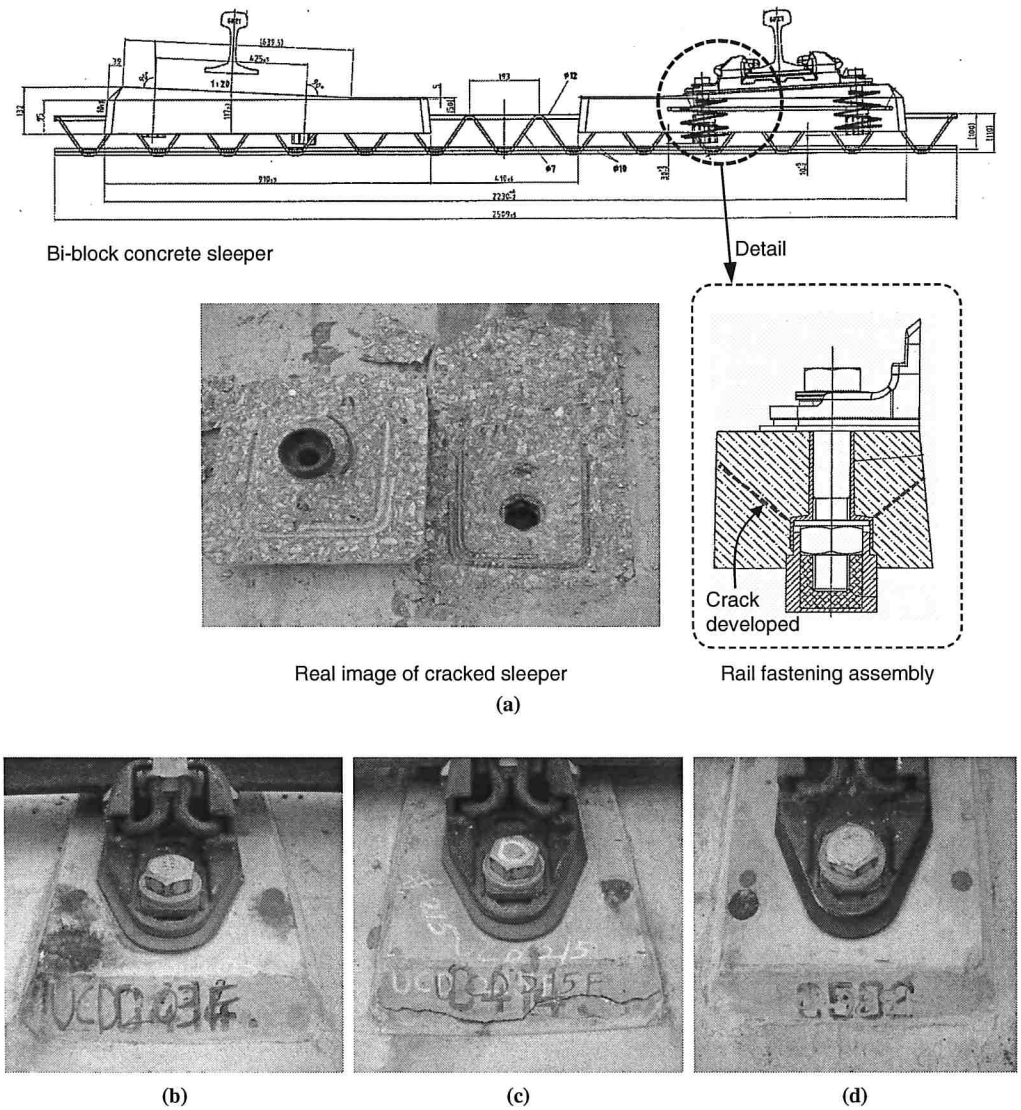


FIGURE 1 Concrete sleeper and fastening assembly used for high-speed railways and selected concrete sleepers tested for internal damage: (*a*) sleeper and fastening assembly, (*b*) sleeper with no cracks, (*c*) sleeper with surface-opening cracks, and (*d*) sleeper with hairline crack and possible latent defects compared with photograph of a cracked sleeper.

depth to the internal anomalies or defects. In impact–echo testing, the thickness of the structure, h, is determined by Equation 1 when the P-wave velocity, v_P, and the resonance frequency, f_R, are available.

$$h = \frac{\beta v_P}{2 f_R} \tag{1}$$

where β is the shape factor and a typical value of β is 0.96 for a concrete pavement slab.

The impact–echo method was applied to detect flaws in reinforced concrete beams and columns (*5*), and the method was later adopted to determine the crack width of concrete plate specimens (*6*). In the 1990s, Sansalone and colleagues discussed the feasibility of using the impact–echo method to detect embedded flaws and outlined a simple law that the lateral dimension of flaws should be at least one-fourth of the flaw depth to be detected (*1, 7*). One important fact regarding the application of the impact–echo method is that it works better with lateral flaws than with vertical or diagonal flaws. This is because the impact–echo method relies on multiple reflections from internal flaws and lateral flaws reflect more energy than vertical flaws.

Surface-Wave Test

The SASW method is a nonintrusive method used to evaluate the stiffness profile in natural geotechnical sites, pavement systems, and concrete structures by measuring the propagation velocities of surface waves (*3, 8*). The measurement records the particle velocities or accelerations induced by propagating surface waves at two locations. The measured velocity or acceleration time histories are transformed to the frequency domain to determine the phase differences between the two receivers for a series of frequencies. The phase differences are then used to determine the propagation velocities of the surface waves at a series of frequencies, which form a phase–velocity dispersion curve.

The inversion analysis is then used to determine the shear wave velocity profile from the dispersion curve.

Because the SASW method was originally developed for evaluation of the elastic moduli of pavement systems (*9–11*), many researchers have studied the feasibility of using the SASW method to characterize surface and embedded cracks in concrete structures or to evaluate the stiffness and strength of concrete material (*12–15*). Bowen used the SASW method to detect crack damage in beam and column elements according to Rayleigh-wave velocities (*14*). Kalinski investigated intact and cracked concrete structural elements using the SASW method (*15*). Recently, the reliability and accuracy of characterizing platelike concrete structures complicated by cracks and internal anomalies was improved by the enhanced resonance search (ERS) technique (*12*). The ERS technique is a combination of the SASW method and the impact–echo method and takes advantages of the two methods to analyze the data measured.

Because the SASW method assesses surface waves propagating through near-surface material, it is more beneficial than other NDT methods for the investigation of both surface-opening cracks and latent cracks. However, for the identification of cracks, the SASW method must be performed with a high-frequency configuration (typically up to 50,000 Hz), which requires a sophisticated and time-consuming test setup, like receiver coupling with epoxy glue.

FRACTAL Technique

To evaluate the stiffness of concrete slabs resting on a subgrade, Cho et al. developed the FRACTAL technique (*4*). The FRACTAL technique was originally proposed as a means of assessing the potential deflection of concrete tracks by measuring the flexural rigidity of concrete railroad tracks. By the FRACTAL technique, the flexural rigidity of concrete tracks is evaluated by means of antisymmetric Lamb-wave velocities (*4, 16*). For FRACTAL tests, one impact source and two accelerometers are used for measurement. In the case of KTX tracks with a 0.6-m-thick concrete slab, receivers are deployed at intervals of 2 m, and the source is offset from the first receiver by 8 m. Typically, FRACTAL tests are performed continuously along the longitudinal axis of railway tracks, and a measurement array is overlapped with an adjacent array by 1 m. The Lamb-wave velocities from FRACTAL tests are compiled together and displayed in a two-dimensional (2-D) representation of the Lamb-wave velocity and wavelength.

Determination of the Lamb-wave velocity by the FRACTAL technique is based on the phase difference between two receivers, which is the same procedure used for the SASW method. Unlike the SASW method, however, the FRACTAL technique uses only the phase-angle information for the Lamb-wave mode, which corresponds to the region of frequencies below the threshold frequency of mode conversion from the Lamb wave to the Rayleigh wave (*4*). The validity of the FRACTAL technique was verified by several comparisons with SASW tests for a subgrade near concrete tracks and impulse response tests for the same concrete slabs (*4, 16*). For the application of the FRACTAL technique to the concrete sleepers of KTX railway tracks, slight modifications to the measurement configuration, including the frequency span, interreceiver spacing, and hammer location, were made. The details are discussed in the following section.

RESPONSE OF CONCRETE SLEEPERS TO STRESS WAVES IN NONDESTRUCTIVE TESTING

In this section, the results of the impact–echo tests, SASW tests, and FRACTAL tests for selected concrete sleepers are compared to determine the optimal method for identifying hairline cracks or latent cracks in concrete sleepers. The tests were performed at a new KTX railway section in the vicinity of Gyeong-Joo, South Korea. Figures 1*b* to 1*d* show the three sleepers tested in this research: an intact and undamaged sleeper (Sleeper 203), a sleeper with surface-opening cracks (Sleeper 215), and a sleeper with a hairline crack and possible latent defects (Sleeper 988). The measurement configurations for the impact–echo tests, SASW tests, and FRACTAL tests are displayed in Figure 2. The results are summarized in the following sections.

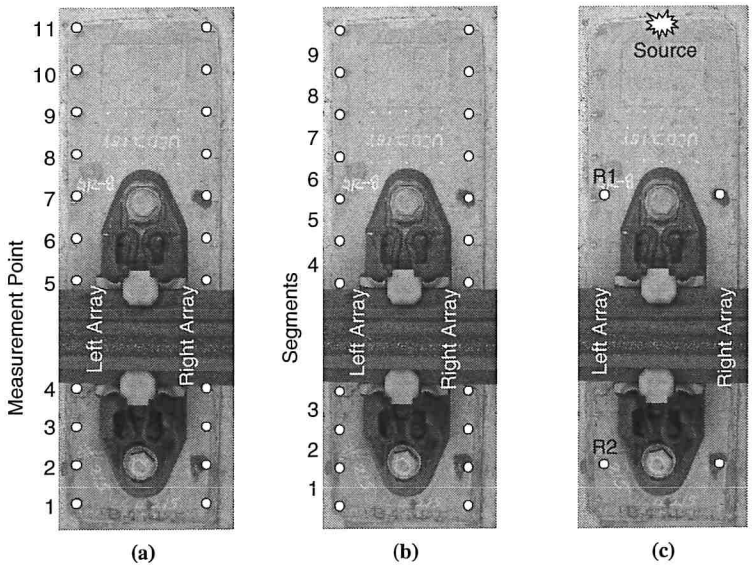


FIGURE 2 Measurement configurations for (*a*) impact–echo tests, (*b*) SASW tests, and (*c*) FRACTAL tests.

Impact–Echo Test

The impact–echo test has been used extensively to identify cracks in platelike structures. For the investigation of latent defects in concrete sleepers, the impact–echo test was the first choice among a variety of NDT techniques. Power spectra are plotted from the impact–echo tests for Sleepers 203, 215, and 988 in Figure 3. As expected, Sleeper 203 (the intact, undamaged sleeper) shows a clear response with a distinct resonance at a frequency of about 17.1 kHz in Figure 3*a*, which corresponds to the sleeper thickness of 13 cm. The measurements were performed at Points 2 and 3, close to the fastening bolt, which are the most plausible locations for an inherent conical crack. Sleeper 215 (the damaged sleeper with surface-opening cracks) had a complicated response with multiple peaks, as shown in Figure 3*b*. The power spectrum with multiple peaks is strong evidence of internal anomalies with multiple occurrences of cracks, delaminations, and voids. Unexpectedly, Sleeper 988 (the sleeper with latent defects) showed a clear response and had a single resonance similar to the one for the intact, undamaged sleeper. The result of the impact–echo test for Sleeper 988 was supposed to show more than two peaks because of the inherent defects and bottom interface. However, apparently there is only one distinct peak, which corresponds to multiple reflections from the bottom interface.

The fact that only one explicit peak was observed from the impact–echo test for Sleeper 988 implies that the latent defect is neither extensively developed nor thick enough to cause multiple reflections. This reasoning can be verified by the 2-D distribution of the impact–echo response measured along the sides of the sleepers. All the contour plots in Figure 4 were constructed from a total of 22 impact–echo tests, which were performed at both the left and the right measurement arrays shown in Figure 2. The variation of the pseudodepth, defined as the P-wave velocity divided by 2 times the frequency, is plotted in Figure 4 with measurement points. Peak locations correspond to the depths of reflection boundaries, such as bottom interfaces or internal anomalies. In the case of Sleeper 203, which has no inherent cracks, Figures 4*a* and 4*b* (the contour plots for the left and the right arrays, respectively) show an increasing depth with an increasing measurement point. This agrees with the physical shape of a sleeper. However, in the case of Sleeper 215, which has surface-opening cracks, no clear peaks were measured at Point 1, 2, or 3 in the left array or at Point 1, 2, 3, or 4 in the right array, as shown in Figures 4*c* and 4*d*. The physical interpretation for no clear peaks is that all the wave energy is scattered away without reflections. That is, wave energy cannot reach the bottom interface, and in turn, no reflections can be made at the bottom interface. On the other hand, Figures 4*e* and 4*f* show a trend similar to the ones shown in Figures 4*a* and 4*b*, which correspond to the intact sleeper without any cracks. That is, the impact–echo tests failed to identify latent defects in Sleeper 988.

The results of the impact–echo tests performed on the sleepers with different internal conditions led to the conclusion that impact–echo tests are good for use only with sleepers with extensive and well-developed cracks and do not provide reliable results for sleepers with latent defects.

SASW Test

SASW tests were also performed with the same sleepers shown in Figure 1 to confirm if the SASW method is valid as a quality assurance tool for the identification of latent cracks in sleepers. The tests were performed at specific segments of the sleepers: Segment 5 of the left array for Sleeper 203, Segment 2 of the left array for Sleeper 215, and Segment 5 of the left array for Sleeper 988. As shown in Figure 2*b*, a total of nine measurements were performed along one side of each sleeper, and both sides were tested by using the same measurement configuration. A total of nine measurement arrays were later deployed side by side to investigate all the material at each side of a sleeper. Interreceiver spacing of 0.075 m was used for the SASW tests, because the target sampling depth in the test was only 0.13 m, which is equal to the thicker side of the sleeper.

The phase spectra of the transfer functions were determined by spectral analysis for signals measured at two accelerometers and are presented in Figure 5. An intact sleeper (Sleeper 203) shows a clear sawtooth pattern trend over a frequency span. The resulting Rayleigh-wave velocities ranged from 2,200 to 2,600 m/s, as plotted in Figure 5*d*. The sleeper with surface-opening cracks (Sleeper 215) revealed a phase spectrum with the significantly jagged shape shown in Figure 5*b*, which leads to a Rayleigh-wave velocity of approximately 200

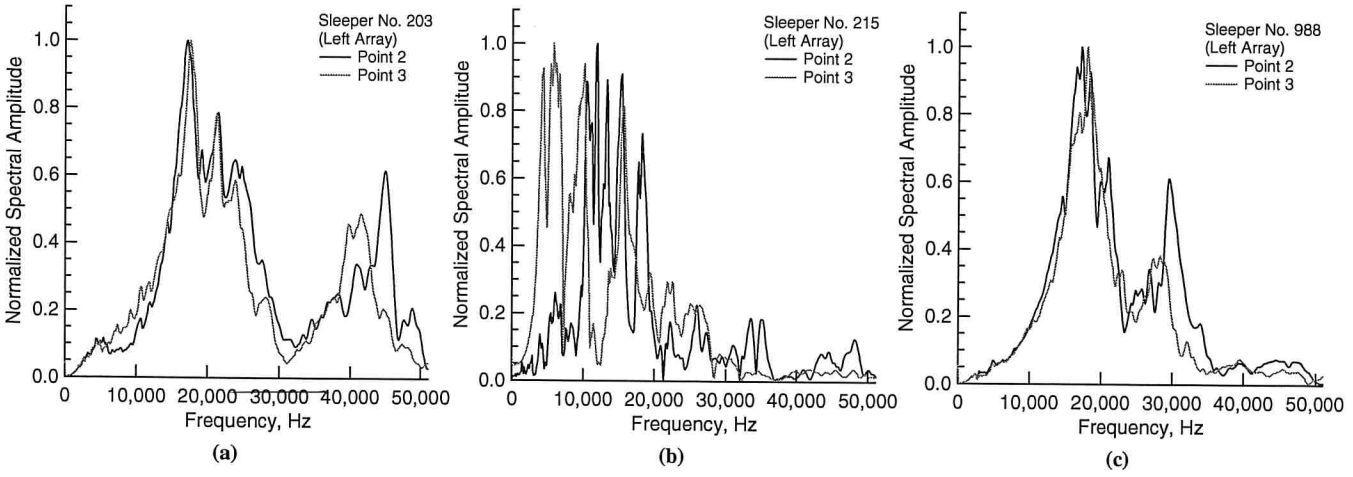


FIGURE 3 Power spectra measured by impact–echo tests for (*a*) sleeper with no cracks, (*b*) sleeper with surface-opening cracks, and (*c*) sleeper with hairline crack and possible latent cracks.

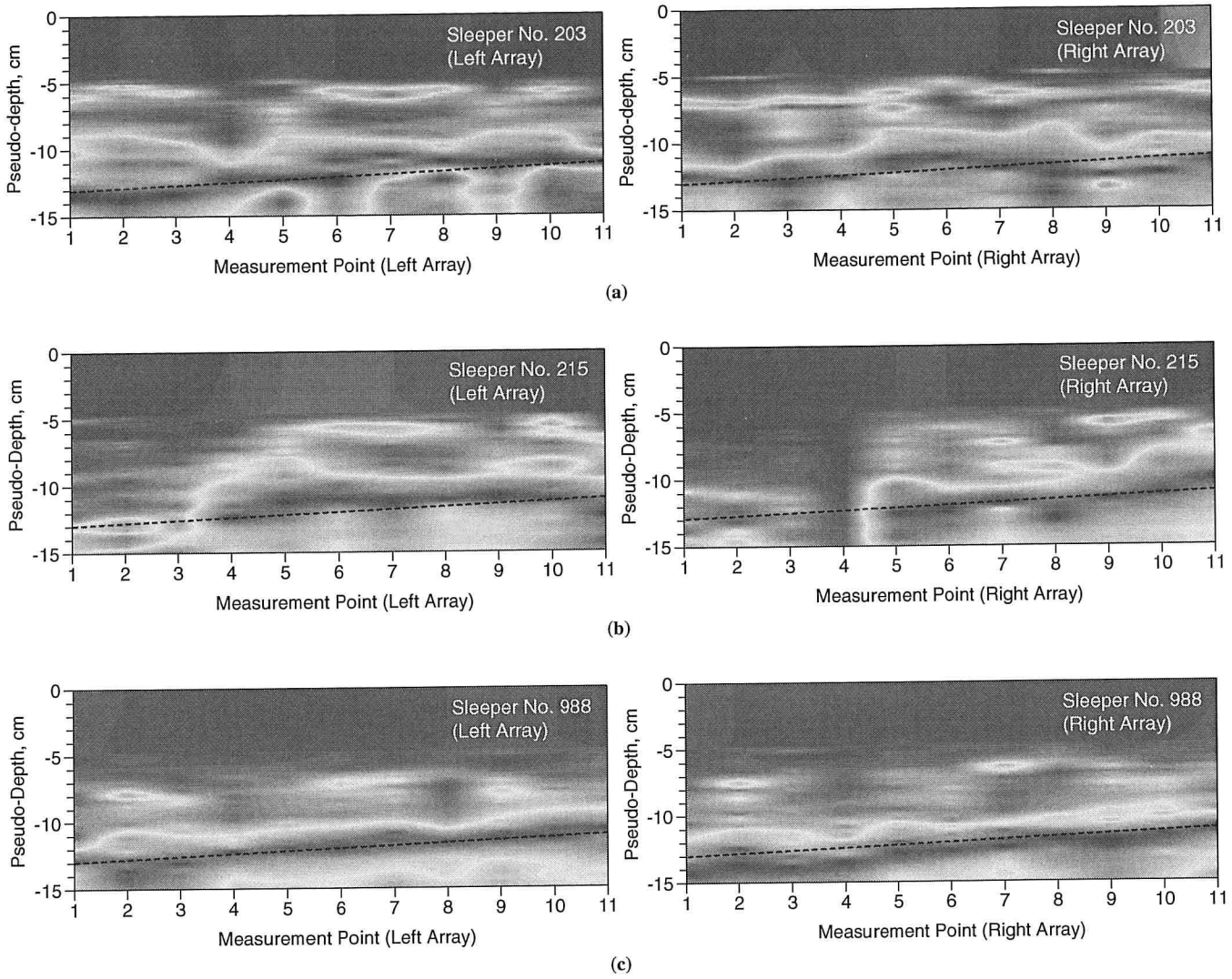


FIGURE 4 2-D representation of power spectra measured by series of impact–echo tests at (*a*) sleeper with no cracks, (*b*) sleeper with surface-opening cracks, and (*c*) sleeper with hairline crack and possible latent cracks.

to 400 m/s. On the other hand, the sawtooth pattern for Sleeper 988 (which has latent cracks) is less prominent than the pattern for the intact sleeper. Specifically, the frequency range between 15,000 and 30,000 Hz is obscure, and the phase spectrum is misleading in its interpretation. The Rayleigh-wave velocities for these three sleepers are compared in Figure 5*d*. The Rayleigh-wave velocities of the sleepers (which are strongly correlated with material stiffness) indicate that Sleeper 203 is the stiffest, Sleeper 215 is the weakest, and Sleeper 988 is a little bit less stiff than the intact sleeper.

All the Rayleigh-wave velocity dispersion curves were combined to plot the 2-D distribution of the Rayleigh-wave velocity in Figure 6. Sleeper 203 shows a uniform distribution of Rayleigh-wave velocities along the segments, whereas Sleeper 215 shows a pronounced contrast between Segments 1 to 3 and Segments 4 to 8. The visualization of the Rayleigh-wave velocity clearly reveals that the material in Segments 1 to 3 is not in good condition. Part of Sleeper 988 appears to have defects. Segments 5, 7, 8, and 9 in the left array and Segments 3, 4, 6, 7, 8, and 9 in the right array show signs of embedded defects and weak material of poor quality.

FRACTAL Test

The FRACTAL technique, which is used to measure the Lamb-wave velocity as an indicator of flexural rigidity in a plate, was also used to investigate the integrity of the concrete sleepers shown in Figure 1. The feature of evaluating the overall flexural rigidity of a sleeper requires the use of only one measurement array on each side of the sleeper shown in Figure 2*c*. FRACTAL tests proceeded by striking the far end of a sleeper and recording the particle accelerations near the fastening bolts. Unlike the original FRACTAL test for a concrete slab track, both the interreceiver spacing and the source offset from the first receiver were set equal to 0.425 m. The frequency span was set equal to 2,000 Hz, whereas that for a concrete slab track is 100 Hz. The results of the FRACTAL tests for Sleepers 203, 215, and 988 are summarized in a phase-spectrum format with the corresponding Lamb-wave velocity dispersion curve, as shown in Figure 7.

The principle of the FRACTAL technique, which is that the flexural rigidity of plates (in this study, the sleepers) can be evaluated by means of Lamb-wave velocities, is apparent in Figure 7. Sleeper 215,

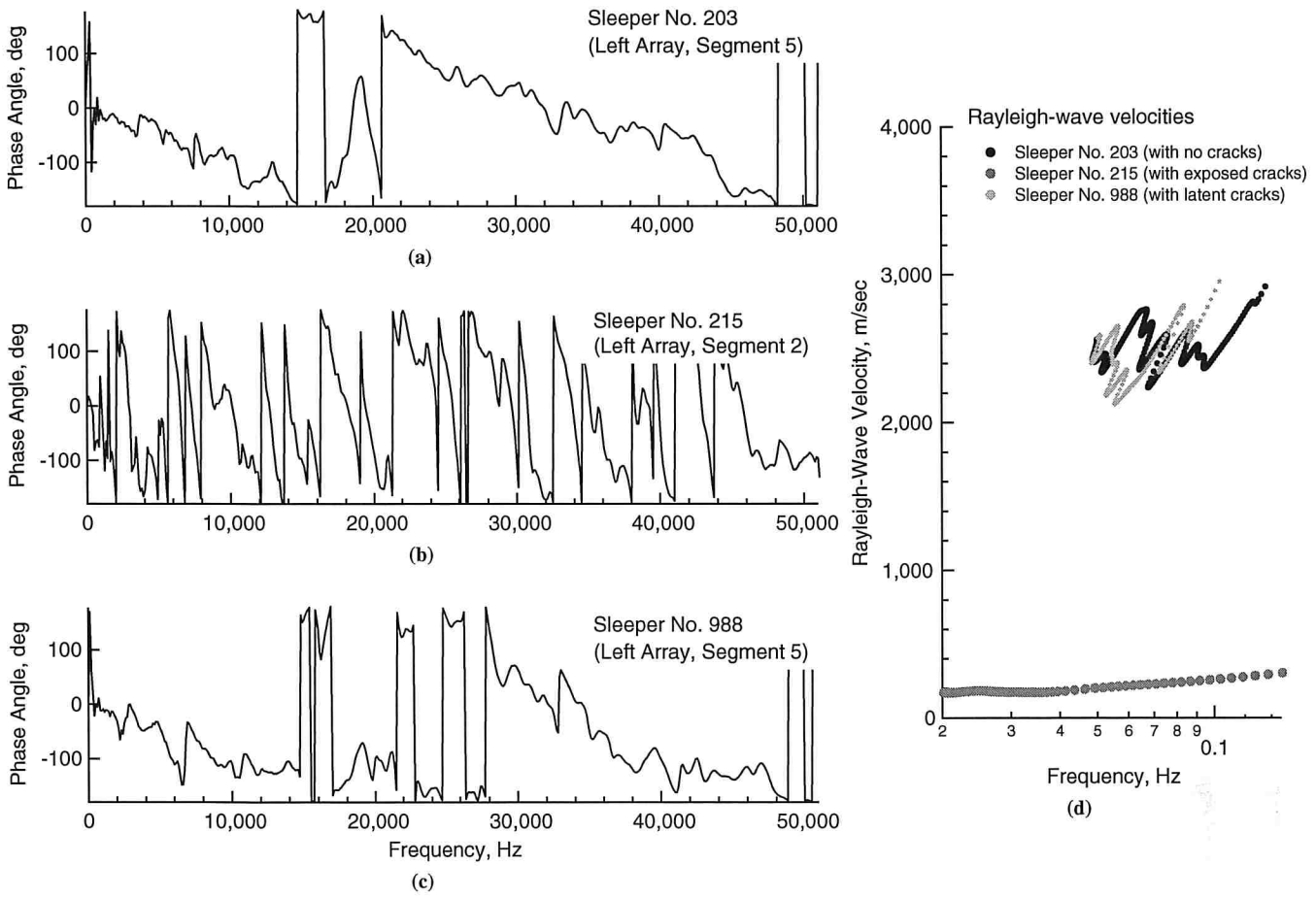


FIGURE 5 Phase-difference spectra and Rayleigh-wave velocities measured by a series of SASW tests at (*a*) sleeper with no cracks, (*b*) sleeper with surface-opening cracks, and (*c*) sleeper with a hairline crack and possible latent cracks, and (*d*) Rayleigh-wave velocity dispersion curve.

which has surface-opening cracks, has a Lamb-wave velocity of less than 600 m/s, whereas Sleeper 203, which has no cracks, has Lamb-wave velocities as high as 2,000 m/s. That is, a Lamb-wave velocity of less than about 600 m/s in a short wavelength range may indicate a sleeper with a defect. Unlike impact–echo tests and SASW tests, a single FRACTAL test should be sufficient to evaluate the quality of a sleeper, that is, to examine whether any cracks are embedded in a sleeper, because flexural rigidity or Lamb-wave velocity is an integrated parameter for the structural integrity of a sleeper.

A practical example of the use of FRACTAL tests is shown in Figure 8. A total of 33 sleepers were tested by the FRACTAL technique, and the resulting Lamb-wave velocity dispersion curves were combined together to investigate the 2-D distribution of the Lamb-wave velocities. As shown in Figure 8, sleepers with defects are easily identified by the large contrasts in Lamb-wave velocities from those of adjacent sleepers.

OPTIMAL PROCEDURE FOR IDENTIFICATION OF LATENT CRACKS IN SLEEPERS

The previous section provided the results of investigations of the impact–echo, SASW, and FRACTAL test methods for their reliability and feasibility in identifying latent cracks in sleepers. For the investigation, three different sleepers were studied, including a sleeper with no cracks, a sleeper with surface-opening cracks, and a sleeper with latent cracks.

The impact–echo method is reliable only when cracks are widely spread and well developed. The method is not sensitive enough to identify latent cracks in sleepers. Even the 2-D presentation of the responses from a series of impact–echo tests performed along a measurement array was not helpful in locating latent cracks. This is probably because latent cracks are too narrow to reflect most of the stress-wave energy back to the surface. On the other hand, the SASW method is more accurate and reliable for identifying local anomalies embedded in a sleeper, such as latent cracks. However, the time and effort required for SASW tests are significant. In this research, a total of 18 SASW tests were performed for one sleeper, and the measurement time increased to about 2 h. The FRACTAL test is fast and straightforward. Only 1 min should be enough for FRACTAL testing at one sleeper. The reliability and accuracy of the FRACTAL test are also comparable to those of the SASW method.

Finally, the optimal procedure suggested for use for the identification of latent cracks in sleepers is a two-stage procedure. The first stage is to scan sleepers by the FRACTAL technique and sort out the sleepers with potential defects. The second stage is to apply the SASW technique alongside a sleeper to confirm if there are latent defects within. This suggested procedure was applied to

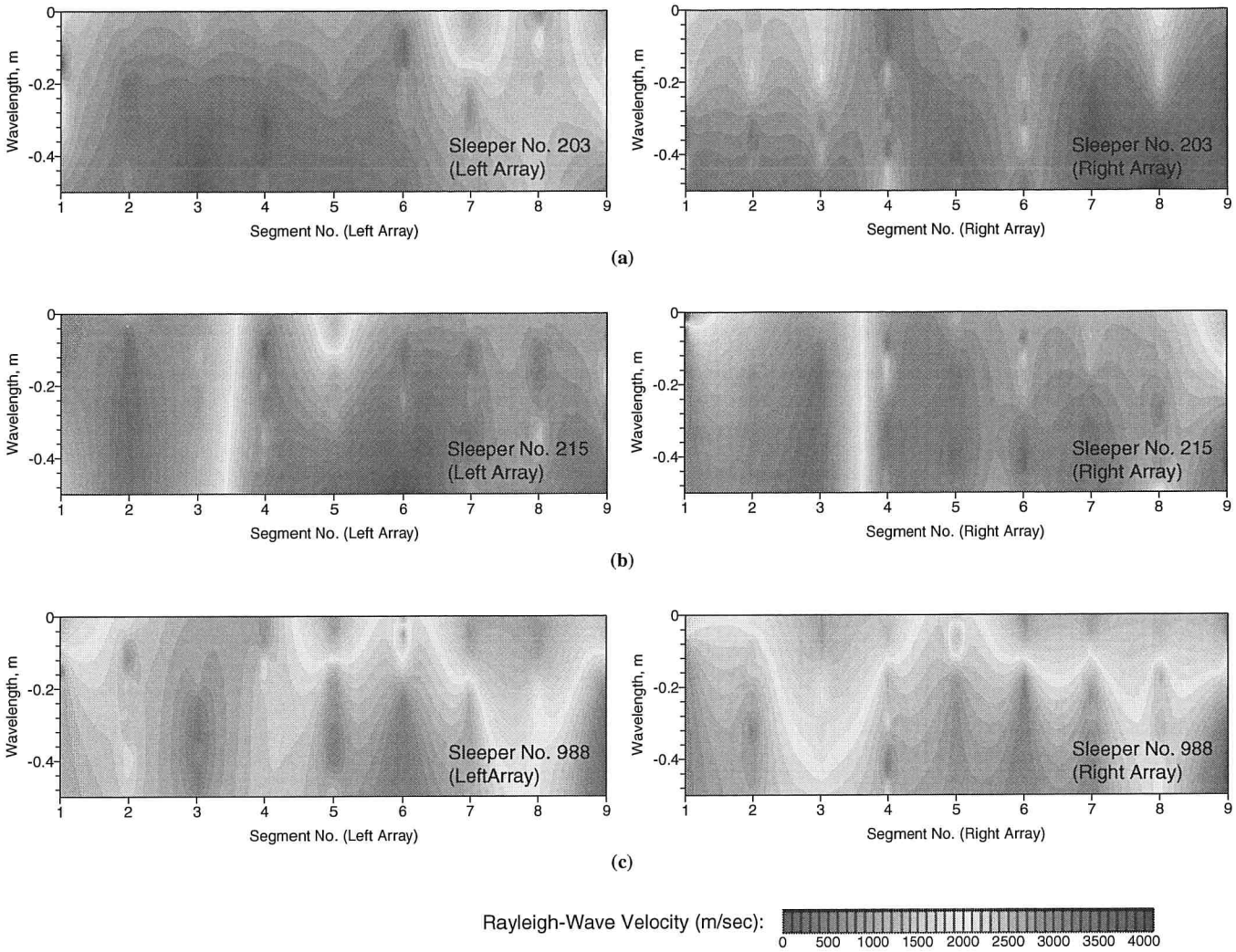


FIGURE 6 2-D representation of Rayleigh-wave velocities measured by a series of SASW tests at (*a*) sleeper with no cracks, (*b*) sleeper with surface-opening cracks, and (*c*) sleeper with a hairline crack and possible latent defects.

18,177 sleepers. The results of those tests are summarized in the following section.

FIELD APPLICATIONS IN CONTEXT OF SOUTH KOREAN HIGH-SPEED RAIL

An intensive investigation of latent cracks in the concrete sleepers of high-speed railways in South Korea was performed by using the nondestructive procedure proposed in this research. A total of 18,177 sleepers were investigated in 4 months. The investigation was conducted in two steps: a preliminary investigation was conducted by use of the FRACTAL technique, and this was followed by an in-depth investigation by use of the SASW method, as shown in Figures 9*a* and 9*b*. The typical time requirements for one set of FRACTAL and SASW tests were about 1 min and 2 h, respectively. The more time demands for SASW measurements resulted from the larger number of test arrays required and the use of a sophisticated coupling method for the accelerometers. One FRACTAL test includes two arrays for one sleeper, whereas one SASW test should include 18 arrays for one sleeper. For expedience, a dedicated hardware system and specialized software were used for the FRACTAL and SASW tests. These are shown in Figures 9*c* and 9*d*, respectively. The measurement system enabled the testing of about 500 sleepers by FRACTAL tests and four sleepers by SASW tests per day.

To determine if a sleeper is intact, without hairline cracks or latent cracks, decision criteria for the Lamb-wave velocity had to be established. To establish the decision criteria, a preliminary investigation consisting of 76 FRACTAL tests and 61 SASW tests was performed only with intact sleepers. The resulting averages and standard deviations for the tests are provided in Figure 10. A simple statistical calculation provides a 95% confidence level of the Lamb-wave velocity and the Rayleigh-wave velocity for a good sleeper. The lower limits of the Lamb-wave and Rayleigh-wave velocities for a good sleeper without latent cracks at a 95% confidence level are 1,445 and 2,505 m/s, respectively.

In this study, some of the 18,177 sleepers tested were selected to validate the FRACTAL technique as an optimal method for assessment of the structural integrity of concrete sleepers. First, three groups of sleepers from construction sections, Construction Sections 11-1, 11-2, and 12-1, were tested by the FRACTAL technique; and defects were found in some of the sleepers, as shown

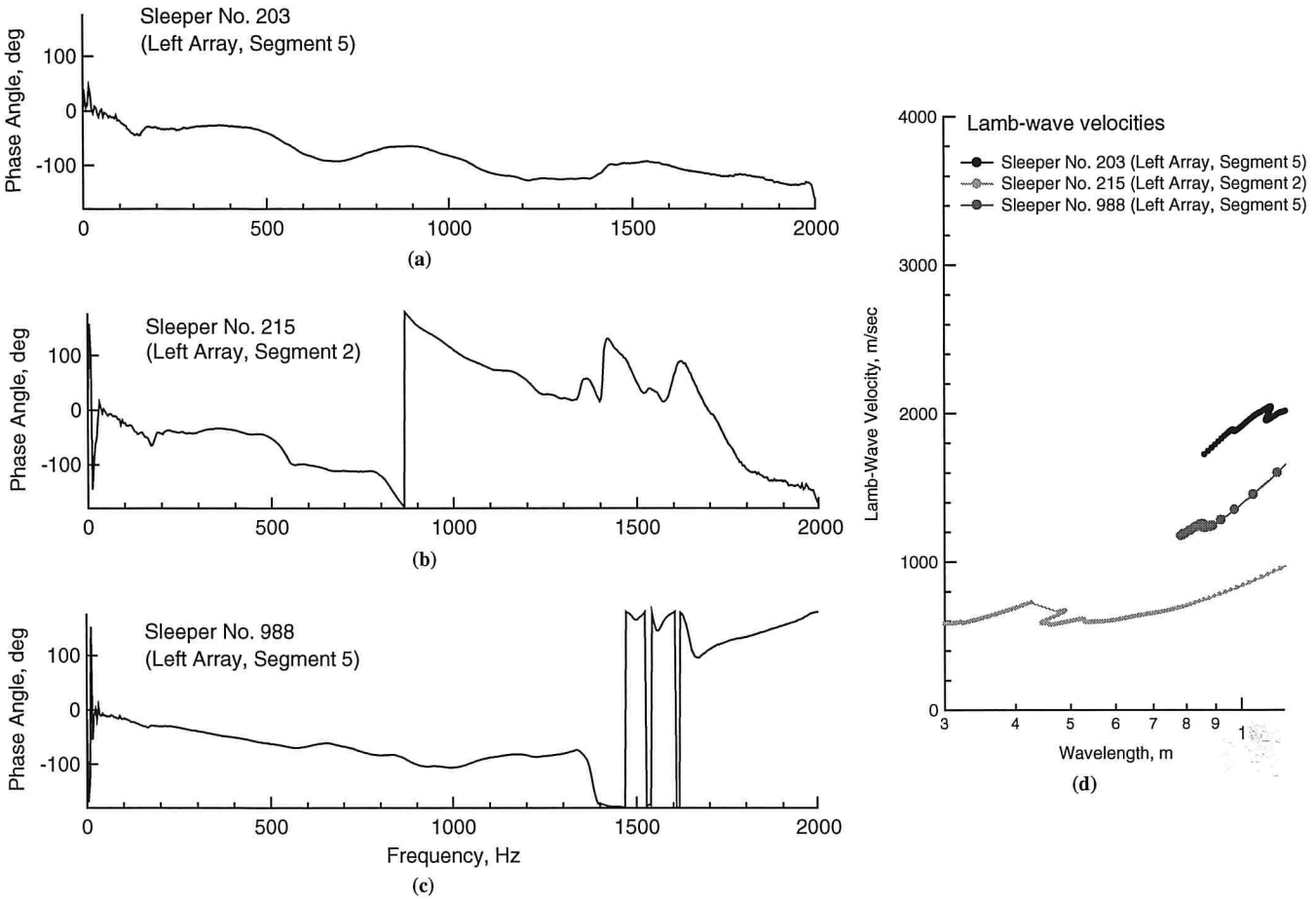


FIGURE 7 Phase-difference spectra and phase velocities for Lamb waves measured by FRACTAL tests at (*a*) sleeper with no cracks, (*b*) sleeper with surface-opening cracks, (*c*) sleeper with a hairline crack and possible latent cracks, and (*d*) Lamb-wave velocity dispersion curve.

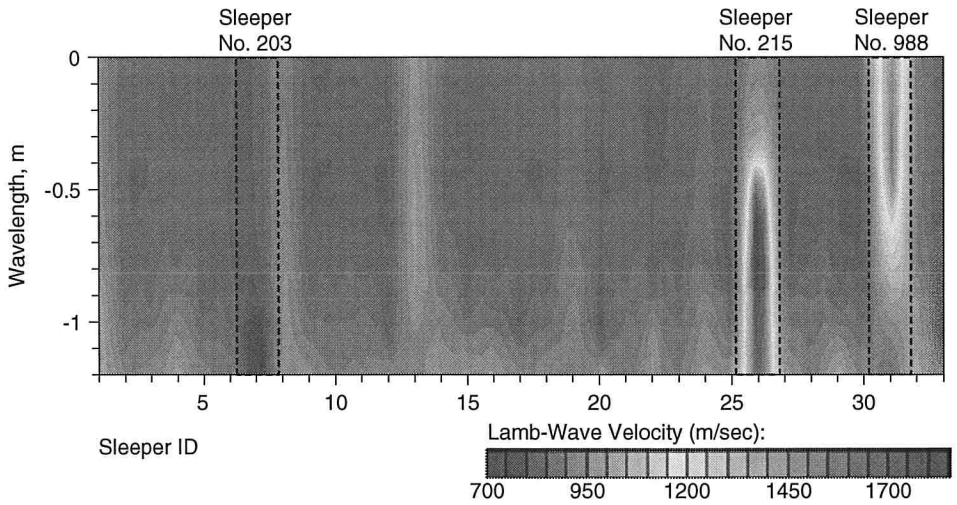


FIGURE 8 2-D representation of Lamb-wave velocities measured by series of FRACTAL tests at total of 33 sleepers (Sleeper ID = sleeper identifier).

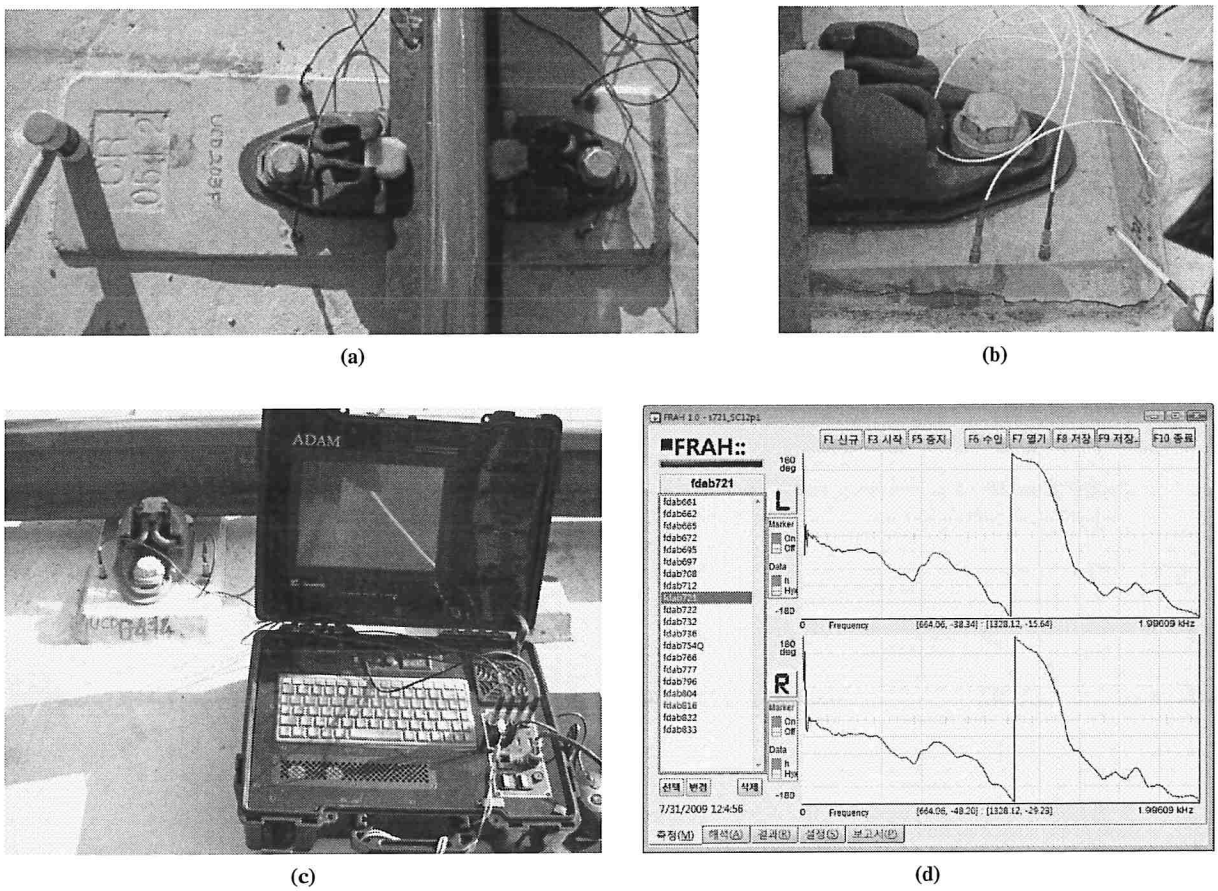


FIGURE 9 Field applications of FRACTAL technique to concrete sleepers of South Korean high-speed railways: (*a*) test setup for FRACTAL technique, (*b*) test setup for SASW method, (*c*) data acquisition hardware, and (*d*) software used for FRACTAL and SASW tests.

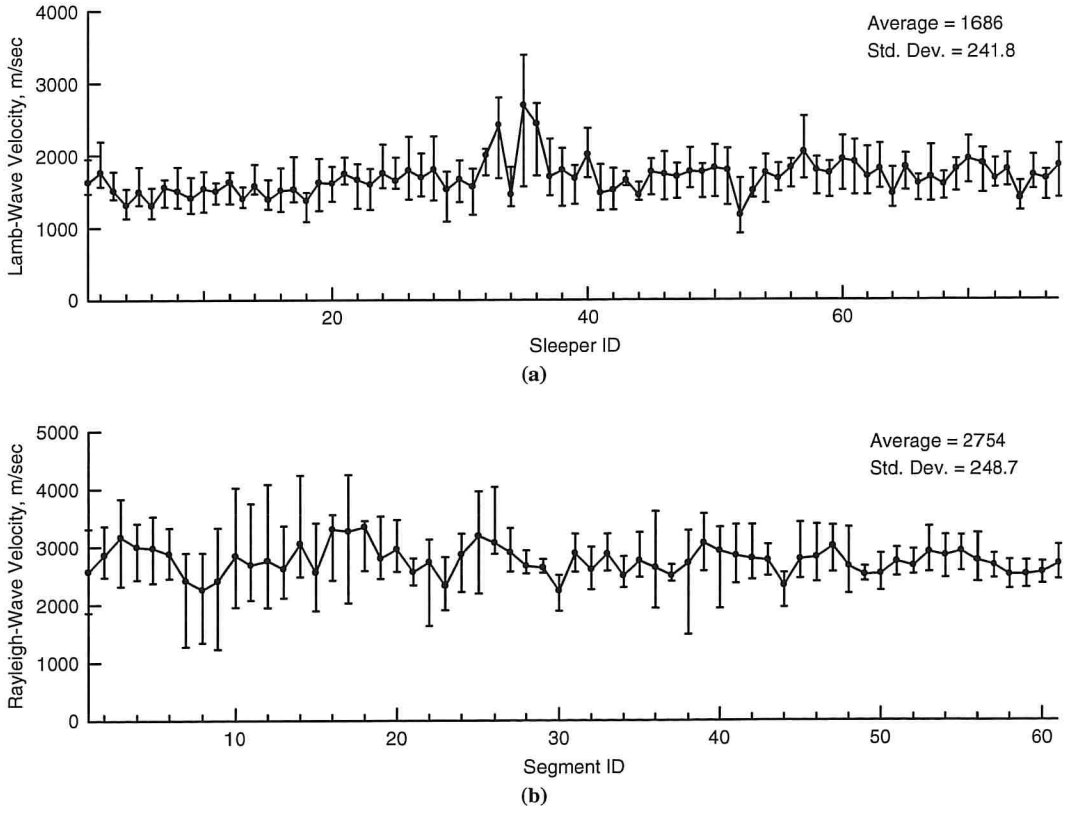


FIGURE 10 Average and standard deviation (Std. Dev.) of Lamb-wave and Rayleigh-wave velocities measured by FRACTAL and SASW tests for sleepers of high quality: (*a*) Lamb-wave velocity variation and (*b*) Rayleigh-wave velocity variation.

in Figure 11. The suspicious sleepers have Lamb-wave velocities significantly less than 1,445 m/s. SASW tests then confirmed that the sleepers are defective, as determined from the Rayleigh-wave velocity profile. In the case of Construction Section 11-1, three suspicious sleepers were found. All of the suspicious sleepers were found to have Lamb-wave velocities of less than 1,000 m/s. On the basis of the Rayleigh-wave velocity criteria (2,505 m/s), intensive SASW tests revealed that there are latent defects at Segments 2 and 6 in the left array and at Segments 3 and 7 in the right array. Likewise, all of the other suspicious sleepers detected by FRACTAL tests were confirmed to have latent defects by intensive SASW tests, as shown in Figures 12*c* to 12*f*.

In the case of Sleeper 412, the sleeper is apparently intact and has no surface-opening cracks. However, during surface drying after a slight rainfall on Sleeper 412, a hairline crack was discovered at Segment 2 in the right array. Interestingly, both the FRACTAL test and the SASW tests also indicated that Sleeper 412 had defects. This supports the reliability and the validity of the FRACTAL technique for identifying damaged sleepers.

SUMMARY AND CONCLUSIONS

This paper describes the implementation of a technique for identifying sleepers internally damaged by the expansion of freezing water. This technique is based on the flexural rigidity of a sleeper rather than its material stiffness or the depth of internal flaws. To verify the validity and reliability of the proposed technique, extensive research was performed for sleepers in three different conditions: a sleeper with no cracks, a sleeper with surface-opening cracks, and a sleeper

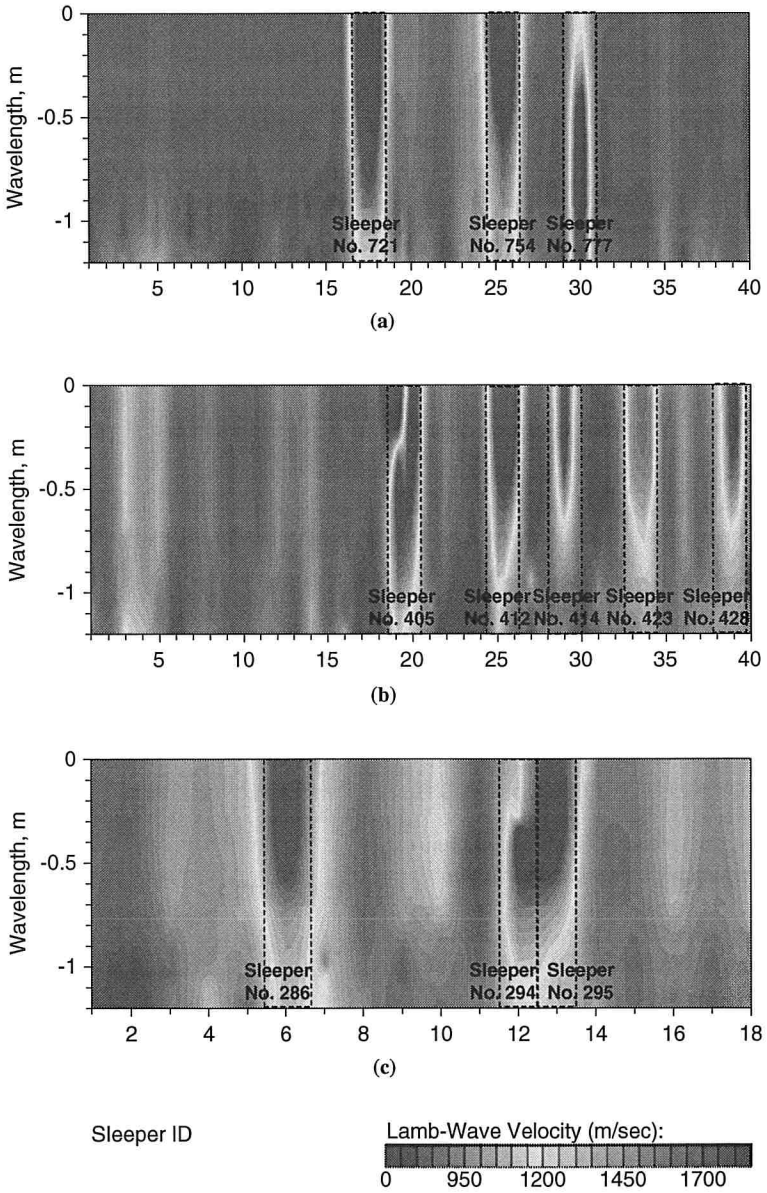


FIGURE 11 2-D representation of Lamb-wave velocities measured by series of FRACTAL tests at different sets of sleepers: (*a*) Construction Section 12-1, (*b*) Construction Section 11-1, and (*c*) Construction Section 11-2.

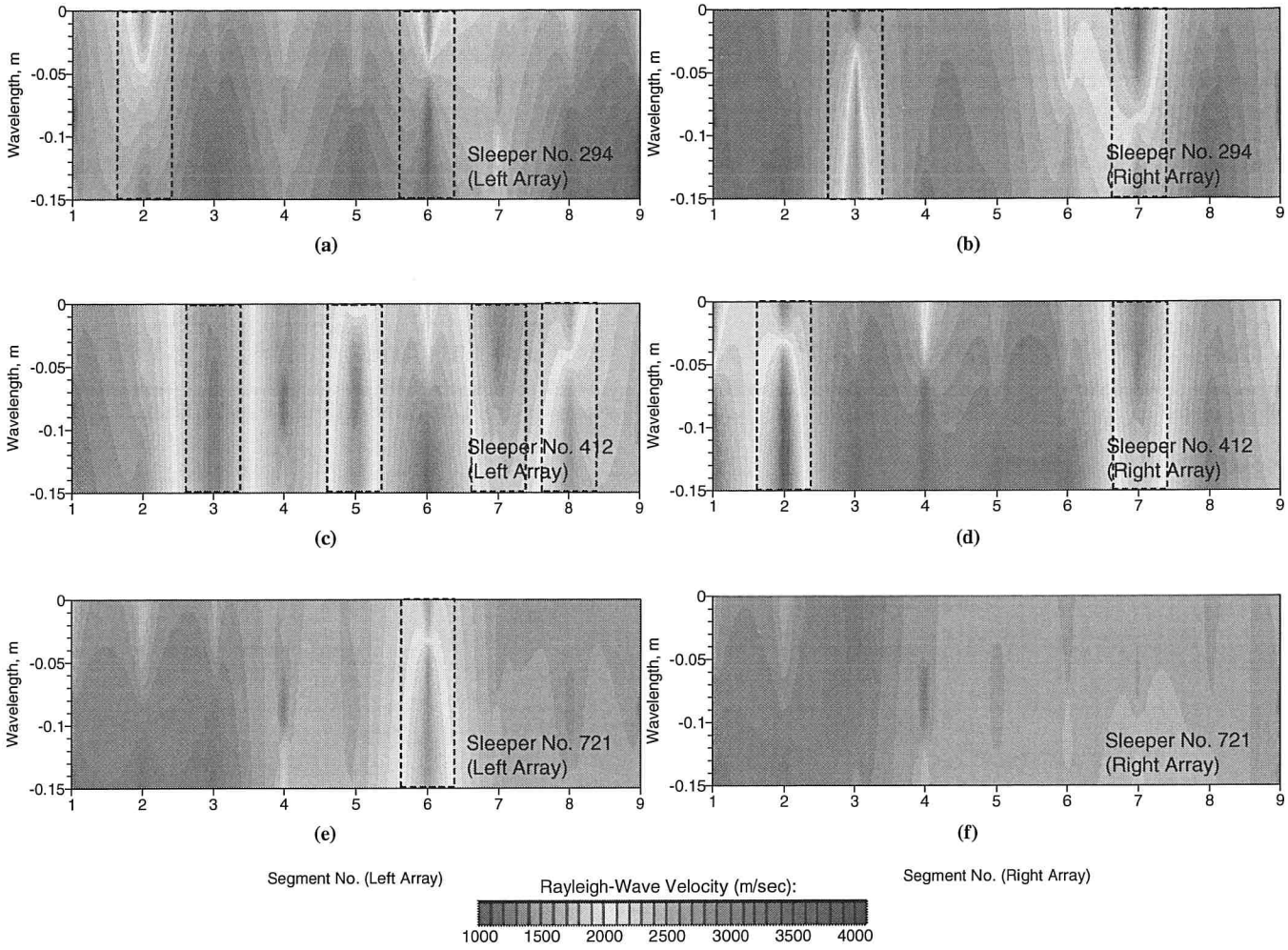


FIGURE 12 2-D profile of Rayleigh-wave velocities measured by SASW tests at sleepers with suspicious latent cracks: (*a*) left array of Sleeper 294, (*b*) right array of Sleeper 294, (*c*) left array of Sleeper 412, (*d*) right array of Sleeper 412, (*e*) left array of Sleeper 721, and (*f*) right array of Sleeper 721.

with a hairline crack and possible latent cracks. The findings are as follows:

- The FRACTAL technique was found to be a quick and valid method for identifying sleepers with surface-opening and latent cracks. One minute was a sufficient amount of time to test a sleeper reliably by the FRACTAL technique.
- The SASW method was found to be slower, but it was also reliable and was the only method that identified the locations of internal damage in sleepers.
- The impact–echo method was successful at identifying sleepers with surface-opening cracks, but it was unsuccessful at identifying sleepers with hairline cracks and possible latent cracks.

The results of the in-depth investigation of NDT tests allowed the design of an optimal procedure to test a total of 18,177 sleepers over a limited period of time (4 months). Preliminary analyses of an intact sleeper without any defects were performed to identify decision criteria. With the decision criteria incorporated in the FRACTAL tests, sleepers with hairline cracks or possible latent cracks were successfully identified, and the findings of the FRACTAL tests were found to agree with the results of SASW tests.

ACKNOWLEDGMENTS

This research is part of the project, Development of Future Technology for Railroad (Next Generation High-Speed Railway), supported by the South Korea Ministry of Land, Transport, and Maritime Affairs, and the 2009 Excellent Researcher Grant from Chung-Ang University. Their financial support is gratefully acknowledged.

REFERENCES

1. Sansalone, M., and W. B. Streett. *Impact Echo: Nondestructive Evaluation of Concrete and Masonry.* Bullbrier Press, Ithaca, N.Y., 1997, p. 339.
2. Sansalone, M., and N. J. Carino. *Impact Echo: A Method for Flaw Detection in Concrete Using Transient Stress Waves.* Report NBSIR 86-3452. National Technical Information Service, Springfield, Va., Sept. 1986.
3. Stokoe, K. H., II, S. G. Wright, J. A. Bay, and J. M. Roesset. Characterization of Geotechnical Sites by SASW Method. In *Geophysical Characteristics of Sites,* International Society for Soil Mechanics and Foundation Engineering Technical Committee 10 for the XIII International Conference on Soil Mechanics and Foundation Engineering, International Science Publishers, New York, 1994, p. 1525.
4. Cho, M.-R., S.-H. Joh, and I. H. Lee. Development of a Nondestructive Seismic Technique for Flexural Rigidity of Concrete Track. *KSCE Journal of Civil Engineering,* Vol. 28, No. 6D, 2008, pp. 905–913.

5. Lin, Y., and M. Sansalone. Detecting Flaws in Concrete Beams and Columns Using the Impact–Echo Method. *Materials Journal,* Vol. 89, 1992, pp. 394–405.
6. Cheng, C.-C., and M. Sansalone. Determining the Minimum Crack Width That Can Be Detected Using the Impact–Echo Method. Part 1. Experimental Study. In *Materials and Structures,* Springer Netherlands, 1995, pp. 74–82.
7. Sansalone, M., J.-M. Lin, and W. B. Streett. Determining the Depth of Surface-Opening Cracks Using Impact-Generated Stress Waves and Time-of-Flight Techniques. *Materials Journal,* Vol. 95, 1998, pp. 168–177.
8. Joh, S.-H. *Advances in Interpretation and Analysis Techniques for Spectral-Analysis-of-Surface-Waves (SASW) Method.* PhD dissertation. University of Texas at Austin, 1996.
9. Heisey, J. S., K. H. Stokoe II, and A. H. Meyer. Moduli of Pavement Systems from Spectral Analysis of Surface Waves. In *Transportation Research Record 852,* TRB, National Research Council, Washington, D.C., 1982, pp. 22–31.
10. Nazarian, S., and K. H. Stokoe II. Nondestructive Testing of Pavements Using Surface Waves. In *Transportation Research Record 993,* TRB, National Research Council, Washington, D.C., 1984, pp. 132–144.
11. Nazarian, S., K. H. Stokoe II, R. C. Briggs, and R. Rogers. Determination of Pavement Layer Thicknesses and Moduli by SASW Method. In *Transportation Research Record 1196,* TRB, National Research Council, Washington, D.C., 1988, pp. 133–150.
12. Kang, T.-H., S.-H. Joh, and M.-R. Cho. Structural Integrity Assessment of Tunnel Concrete Lining by the Resonance Search Technique. *Key Engineering Materials,* Vol. 321–323, 2006, pp. 302–305.
13. Cho, M.-R., S.-H. Joh, S.A. Kwon, and T.-H.Kang. Nondestructive In-Place Strength Profiling of Concrete Pavements by Resonance Search Technique. Presented at 86th Annual Meeting of the Transportation Research Board, Washington, D.C., 2007.
14. Bowen, B. R. *Damage Detection in Concrete Elements with Surface Wave Measurements.* PhD dissertation. Air Force Institute of Technology, Wright–Patterson Air Force Base, Ohio, 1992, p. 207.
15. Kalinski, M. E. *Measurements of Intact and Cracked Concrete Structural Elements by the SASW Method.* Master's thesis. University of Texas at Austin, 1994.
16. Joh, S.-H., I.-W. Lee, S. J. Hwang, T.-H. Kang, and S. C. Kim. Evaluation of Concrete-Track Deformation for High-Speed Railways by Characteristic Stiffness. Presented at 2009 National Conference of Korean Society for Railways, 2009.

The Railway Maintenance Committee peer-reviewed this paper.

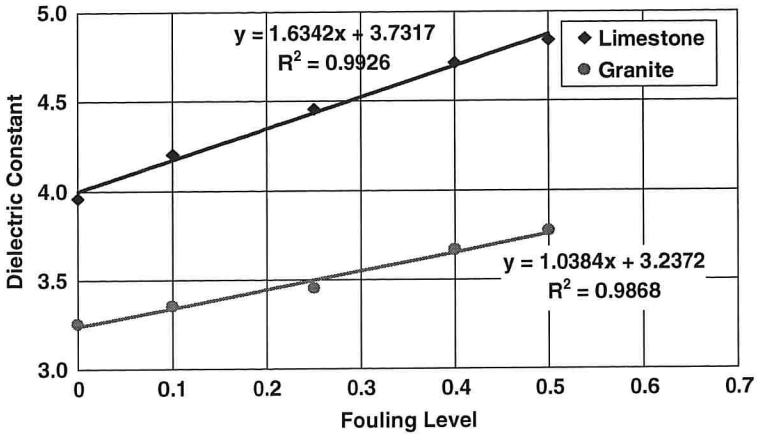


FIGURE 7 Dielectric constants of ballasts fouled by various percentages of dry clay.

in the dielectric constants are from 3.9 to 9.1 and from 4.2 to 10.5 for granite and limestone, respectively.

In addition to the dielectric constant measurement, laboratory testing data were also used to build the STFT color maps for ballasts at various fouling and moisture levels. The STFT color maps for several typical cases are presented in this paper to demonstrate how the STFT method works for ballast fouling and moisture condition assessment.

Figure 9*a* is the STFT spectrum for 36-in.-thick (915-mm-thick) dry clean ballast. It is clear that the frequency energy attenuates gradually and relatively smoothly for clean ballast. However, in Figure 9*b*, which presents the STFT spectrum for 24-in.-thick (610-mm-thick) dry clean ballast on top of 12-in.-thick (305-mm-thick) 50% fouled ballast, a sudden energy drop can be observed at about 26 in. (635 mm). This indicates that the 11-in.-thick (279-mm-thick) ballast at the bottom is fouled. The dry clay, which was applied to fill 50% of the 12-in.-thick (305-mm-thick) clean ballast air void volume, actually filled 10 in. (254 mm) instead of 100% filling of 6 in. (152 mm). This was considered reasonable, as the air voids were 60% filled. For the same ballast shown in Figure 9*b*, when 50% water by the air void volume of the 12-in.-thick (305-mm-thick) ballast was applied, the STFT spectrum in Figure 9*c* was obtained. Compared with the STFT spectrum in Figure 9*b*, a high-energy area (hot color) was observed at the bottom. The high energy is an indication of water accumulation at this location, which resulted in strong signal reflection.

FIELD VALIDATION

To further validate the effectiveness of the STFT method for ballast assessment by the use of GPR data, the field GPR data collected at the Orin Subdivision in eastern Wyoming were analyzed. As shown in Figure 10, the GPR data were collected by using 2-GHz air-coupled antennae mounted on a high-rail vehicle suspended above the rail track. The same data collection system used in the laboratory testing (GSSI SIR-20) was used. To reduce the influence of the rails and ties, the antennae were placed 2 ft (600 mm) away from the rails, which actually collected data from the shoulders of the railroad ballast. Field ballast samples were also collected at 6-in.-deep (152-mm-deep) intervals from the shoulders at selective locations. The aggregate type of the railroad ballast is granite. The moisture content of the collected samples was measured, and aggregate gradation analyses were conducted. The fouling index, F_I, was calculated for each sample to measure the ballast fouling condition. According to Selig and Waters

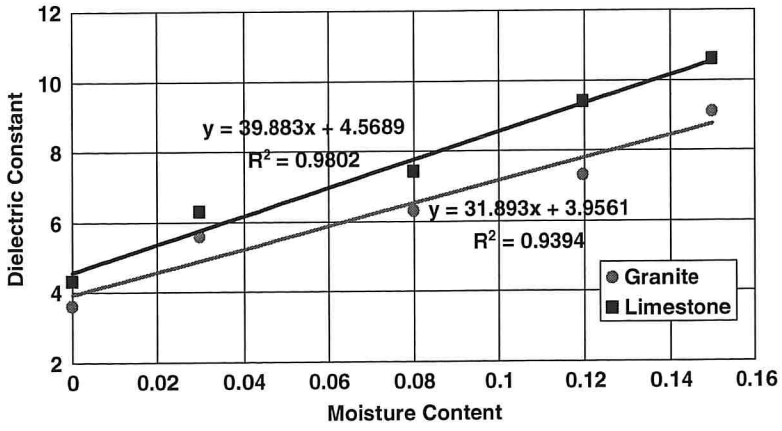


FIGURE 8 Dielectric constants of ballasts at various moisture contents.

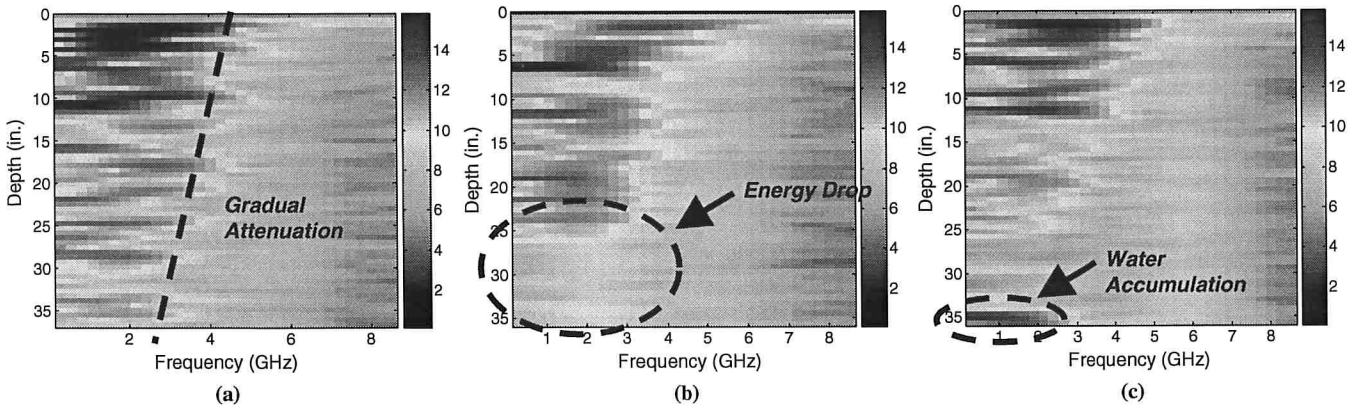


FIGURE 9 STFT spectra of ballast under various fouling and moisture conditions: (*a*) 36-in. clean ballast, (*b*) 24-in. clean ballast on top of 12-in. 50% fouled ballast, and (*c*) ballast in panel *b* with 50% water by air void volume of 12-in. clean ballast.

(*13*), the fouling index can be calculated by using the following equation:

$$F_I = P_4 + P_{200} \tag{5}$$

where P_4 represents the weight percentage of particles passing a No. 4 (4.75-mm) sieve, and P_{200} is the percentage of fine particles passing a No. 200 (0.075-mm) sieve. A value less than 10% is considered clean ballast. If the value is 10% to 20%, the ballast is moderately fouled. If the value is greater than 20%, the ballast is considered seriously fouled. Table 1 shows the fouling indexes for four cases, and Figure 11 presents their corresponding STFT color maps. The thickness of the ballast in the STFT color maps was calculated by using the laboratory-measured dielectric constants, according to their moisture content.

According to Figure 11*a*, a clear energy drop in Case 1 starts at about 25 in. (635 mm) and goes all the way down to 36 in. (915 mm), which matches well with the ground truth data in Table 1: serious ballast fouling occurs at the depth between 24 to 36 in. (610 to 915 mm). It should be noted that because the field ballast samples were collected at 6-in. (152-mm) intervals, a large fouling index could be an indication of fouling at any depth within this interval. For example, in Case 1, the fouling index for samples at the depth of 24 to 30 in. (610 to 762 mm) is 75%. It does not correspond to fouling at 24 in. (610 mm). Instead, the serious fouling could start at any depth between 24 and 30 in. (610 and 762 mm).

According to Figure 11*b*, a clear energy drop in Case 2 starts at about 15 in. (381 mm). However, a strong-energy area is located at a depth of about 20 in (508 mm), and then the energy continues dropping. In the STFT color maps, the strong-energy area corresponds to the strong reflection in the time-domain signal. Table 1 indicates that the serious ballast fouling starts at 12 to 18 in. (305 to 457 mm), which proves that the fouling location indicated by the STFT map is reasonable. The strong-energy area reflects water accumulation at 20 in. (508 mm), which is supported by the fact that there was significant rainfall before the data collection.

In Figure 11*c*, a clear energy drop occurs at 17 in. (432 mm), which is in good agreement with the ground truth data: moderate to serious fouling occurred at 12 to 18 in. (305 to 457 mm).

In Figure 11*d*, a small energy drop can be observed between 11 and 14 in. (279 and 356 mm), and a significant energy drop can be detected at about 18 in. (457 mm). The ground truth data in Table 1 indicate that the ballast material for Case 4 is moderately fouled between 12 and 18 in. (305 and 457 mm) and is seriously fouled after 18 in. (457 mm), which matches the STFT color maps very well.

As has been shown in the several case analyses described above, STFT is an effective method of providing a reasonably accurate prediction of the ballast fouling condition and water accumulation by the use of GPR data, especially when accurate dielectric constants are provided.

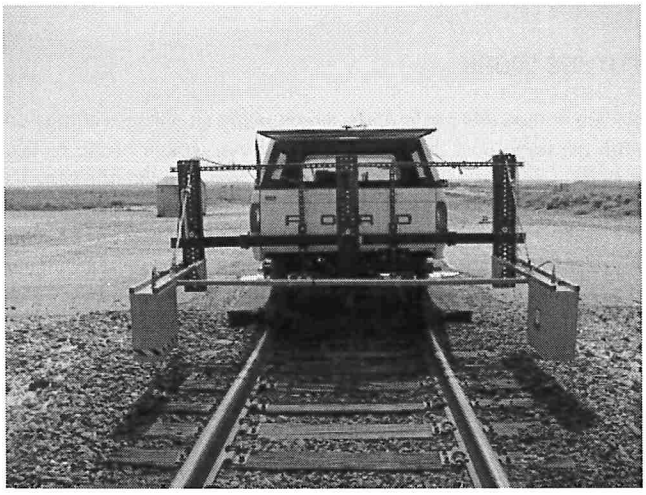


FIGURE 10 GPR equipment on railroad track.

TABLE 1 Fouling Indexes for Cases 1 to 4

	Percent at Depth of					
Case	0–6 in.	6–12 in.	12–18 in.	18–24 in.	24–30 in.	30–36 in.
1	1.7	1.1	2.8	7.8	75	82
2	0	0.2	106	87		
3	9.2	5.9	19.3			
4	5.6	5.0	13.8	34.8		

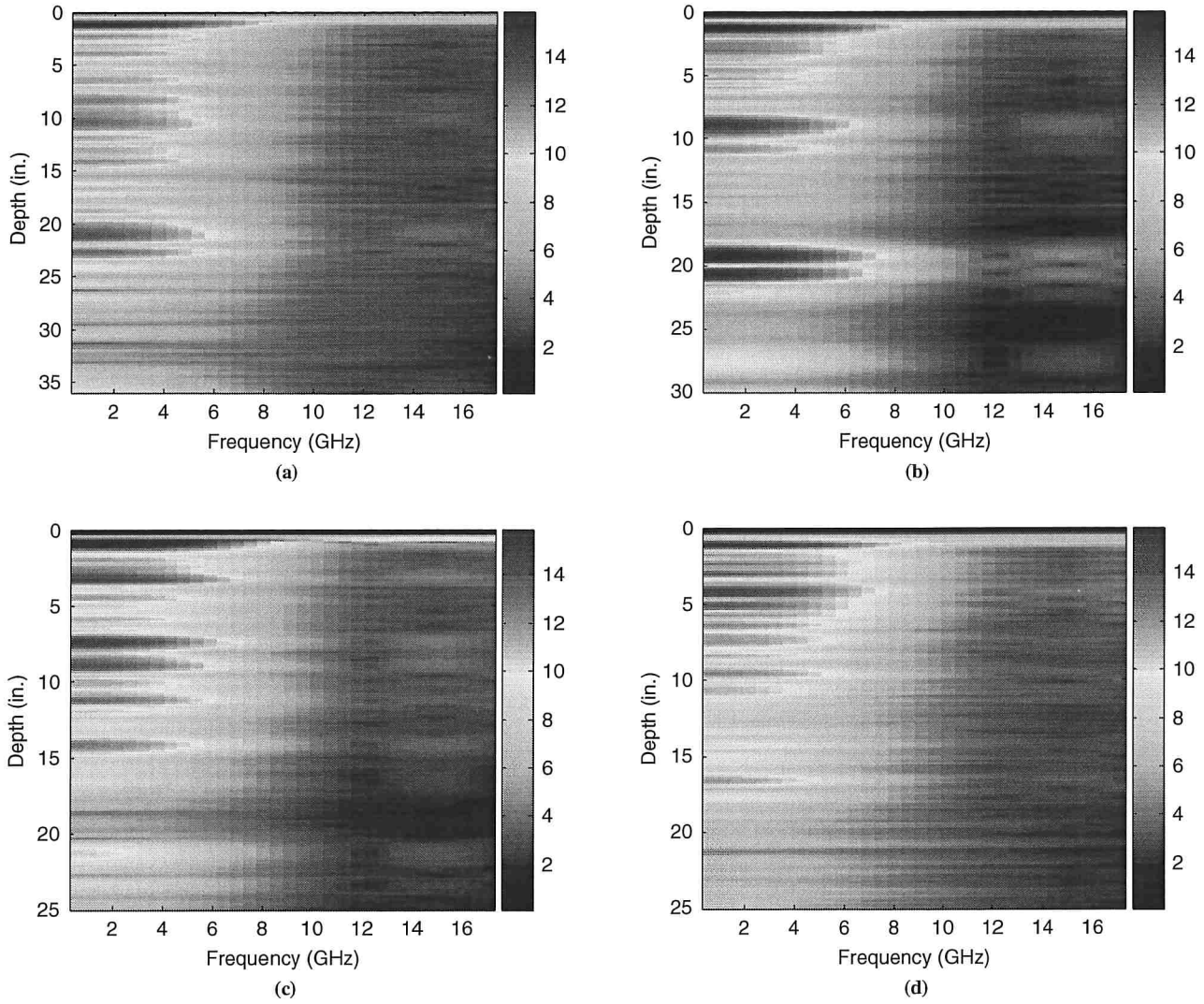


FIGURE 11 STFT color maps for (*a*) Case 1, (*b*) Case 2, (*c*) Case 3, and (*d*) Case 4.

CONCLUSIONS

In this study, controlled laboratory testing was conducted to measure the dielectric constants accurately and build STFT spectra for ballast under various fouling and moisture conditions. The laboratory-measured dielectric constants were then used for the field data analysis to validate the effectiveness of the STFT method for the assessment of ballast fouling. The following summarizes the conclusions of this study:

- Limestone ballast was found to have a greater dielectric constant than granite ballast at the same fouling level.
- At a dry clay fouling level of 0 to 50%, the dielectric constants range from 3.25 to 3.77 and 3.96 to 4.84 for granite and limestone ballast, respectively.
- The dielectric constants of ballast fouled by various percentages of dry clay can be predicted by using linear relationships, as can the dielectric constants of ballast with various moisture contents.
- Moisture can significantly increase the dielectric constant of ballast. A positive linear relationship between the bulk ballast dielectric constant and the moisture content was found.
- The STFT color map can effectively detect locations of fouling and water accumulation if an accurate dielectric constant is used, and this has been validated by both laboratory and field data.

FUTURE WORK

Although the GPR method has shown many advantages compared with the traditional drilling method for assessment of the ballast condition, it is important to recognize its limits so that the benefits of this technique will not be oversold. Detecting ballast fouling by the STFT method is based on the assumption that the GPR data collected are clean and without significant noise. However, during field data collection, many sources of noise (e.g., rails, ties, and radio signals) are generated, and this noise may mask the ballast information in the GPR data. Therefore, although GPR equipment should be properly set up during data collection, efforts are needed to further improve the GPR data filtering and data analysis techniques. To maximize the outcome of assessing ballast fouling by GPR, appropriate guidelines must be developed to assist railroad engineers with the effective use of this nondestructive testing tool.

ACKNOWLEDGMENTS

This research is sponsored by the University of Illinois at Urbana–Champaign (UIUC)–Association of American Railroad (AAR) Technology Scanning Research. The invaluable input of Roger Roberts and Erol Tutumluer is greatly appreciated. The authors also acknowledge the assistance of James Meister and the research project panel for their help and feedback during the study.

REFERENCES

1. Jack, R., and P. Jackson. Imaging Attributes of Railway Track Formation and Ballast Using Ground Probing Radar. *NDT&E International Journal,* Vol. 32, 1999, pp. 457–462.
2. Olhoeft, G. R., and E. T. Selig. Ground Penetrating Radar Evaluation of Railroad Track Substructure Conditions. *Proc., 9th International Conference on Ground Penetrating Radar,* Santa Barbara, Calif., 2002, pp. 48–53.
3. Sussmann, T. R. *Application of Ground Penetrating Radar to Railway Track Substructure Maintenance Management.* PhD thesis. University of Massachusetts, Amherst, 1999.
4. Clark, M. R., R. Gillespie, T. Kemp, D. M. McCann, and M. C. Forde. Electromagnetic Properties of Railway Ballast. *NDT&E International Journal,* Vol. 34, No. 5, 2001, pp. 305–311.
5. Silvast, M., M. Levomaki, A. Nurmikolu, and J. Noukka. NDT Techniques in Railway Structure Analysis. *Proc., World Congress on Railway Research Conference,* Montreal, Quebec, Canada, 2006.
6. Roberts, R., I. L. Al-Qadi, E. Tutumluer, J. Boyle, and T. Sussmann. Advances in Railroad Ballast Evaluation Using 2GHz Horn Antenna. *Proc., 11th International Conference on Ground Penetrating Radar* (CD), Columbus, Ohio, 2006.
7. Al-Qadi, I. L., W. Xie, and R. Roberts. Scattering Analysis of Railroad Ballast Using Ground Penetrating Radar. *Journal of Nondestructive Testing and Evaluation,* Vol. 41, No. 6, 2008, pp. 441–447.
8. Al-Qadi, I. L., W. Xie, R. Roberts, and Z. Leng. Data Analysis Techniques for GPR Used for Assessing Railroad Ballast in High Radio-Frequency Environment. *Journal of Transportation Engineering,* Vol. 136, 2010, p. 392–399.
9. Eriksen, A., J. G. Zetica, and W. Al-Nuaimy. Improved Productivity and Reliability of Ballast Inspection Using Road-Rail Multi-Channel GPR. In *Railway Engineering,* Commonwealth Institute, London, 2004.
10. Lahouar, S. *Development of Data Analysis Algorithms for Interpretation of Ground Penetrating Radar Data.* PhD thesis. Virginia Polytechnic Institute and State University, Blacksburg, 2003.
11. Xie, W., I. L. Al-Qadi, D. L. Jones, and R. L. Roberts. Development of a Time-Frequency Approach to Quantify Railroad Ballast Fouling Condition Using UWB GPR Data. Presented at 87th Annual Meeting of the Transportation Research Board, Washington, D.C., 2008.
12. Oppenheim, A. V., R. W. Schafer, and J. R. Buck. *Discrete-Time Signal Processing.* Pearson Prentice-Hall, Englewood Cliffs, N.J., 2005.
13. Selig, E. T., and J. M. Waters. *Track Geotechnology and Substructure Management.* Thomas Telford Ltd., London, 1994.

The contents of this paper reflect the view of the authors, who are responsible for the facts and the accuracy of the data. The contents do not necessarily reflect the official views or policies of AAR or the UIUC–AAR Program. This paper does not constitute a standard, specification, or regulation.

The Railway Maintenance Committee peer-reviewed this paper.

TRANSPORTATION RESEARCH RECORD:
JOURNAL OF THE TRANSPORTATION RESEARCH BOARD

Peer Review Process

The ***Transportation Research Record: Journal of the Transportation Research Board*** publishes approximately 25% of the more than 3,000 papers that are peer reviewed each year. The mission of the Transportation Research Board (TRB) is to disseminate research results to the transportation community. The Record series contains applied and theoretical research results as well as papers on research implementation.

The TRB peer review process for the publication of papers allows a minimum of 30 days for initial review and 60 days for rereview, to ensure that only the highest-quality papers are published. At least three reviews must support a committee's recommendation for publication. The process also allows for scholarly discussion of any paper scheduled for publication, along with an author-prepared closure.

The basic elements of the rigorous peer review of papers submitted to TRB for publication are described below.

Paper Submittal: June 1–August 1

Papers may be submitted to TRB at any time. However, most authors use the TRB web-based electronic submission process available between June 1 and August 1, for publication in the following year's Record series.

Initial Review: August 15–November 15

TRB staff assigns each paper by technical content to a committee that administers the peer review. The committee chair assigns at least three knowledgeable reviewers to each paper. The initial review is completed by mid-September.

By October 1, committee chairs make a preliminary recommendation, placing each paper in one of the following categories:

1. Publish as submitted or with minor revisions,
2. Publish pending author changes and rereview, or
3. Reject for publication.

By late October, TRB communicates the results of the initial review to the corresponding author indicated on the paper submission form. Corresponding authors communicate the information to coauthors. Authors of papers in Category 2 (above) must submit a revised version addressing all reviewer comments and must include a cover letter explaining how the concerns have been addressed.

Rereview: November 20–January 25

The committee chair reviews revised papers in Category 1 (above) to ensure that the changes are made and sends the Category 2 revised papers to the initial reviewers for rereview. After rereview, the chairs make the final recommendation on papers in Categories 1 and 2. If the paper has been revised to the committee's satisfaction, the chair will recommend publication. The chair communicates the results of the rereview to the authors.

Discussions and Closures: February 1–May 15

Discussions may be submitted for papers that will be published. TRB policy is to publish the paper, the discussion, and the author's closure in the same Record.

Many papers considered for publication in the *Transportation Research Record* are also considered for presentation at TRB meetings. Individuals interested in submitting a discussion of any paper presented at a TRB meeting must notify TRB no later than February 1. If the paper has been recommended for publication in the *Transportation Research Record*, the discussion must be submitted to TRB no later than April 15. A copy of this communication is sent to the author and the committee chair.

The committee chair reviews the discussion for appropriateness and asks the author to prepare a closure to be submitted to TRB by May 15. The committee chair reviews the closure for appropriateness. After the committee chair approves both discussion and closure, the paper, the discussion, and the closure are included for publication together in the same Record.

Final Manuscript Submittal: March 15

In early February, TRB requests a final manuscript for publication—to be submitted by March 15—or informs the author that the paper has not been accepted for publication. All accepted papers are published by December 31.

Paper Awards: April to January

The TRB Executive Committee has authorized annual awards sponsored by Groups in the Technical Activities Division for outstanding published papers:

- Charley V. Wootan Award (Policy and Organization Group);
- Pyke Johnson Award (Planning and Environment Group);
- K. B. Woods Award (Design and Construction Group);
- Patricia F. Waller Award (Safety and System Users Group);
- D. Grant Mickle Award (Operations and Preservation Group); and
- John C. Vance Award (Legal Resources Group).

Other Groups also may nominate published papers for any of the awards above. In addition, each Group may present a Fred Burggraf Award to authors 35 years of age or younger.

Peer reviewers are asked to identify papers worthy of award consideration. Each Group reviews all papers nominated for awards and makes a recommendation to TRB by September 1. TRB notifies winners of the awards, which are presented at the following TRB Annual Meeting.

Transportation Research Board
www.TRB.org

TRANSPORTATION RESEARCH BOARD
500 Fifth Street, NW
Washington, DC 20001

ADDRESS SERVICE REQUESTED

NON-PROFIT ORG.
U.S. POSTAGE
PAID
WASHINGTON, D.C.
PERMIT NO. 8970

THE NATIONAL ACADEMIES™
Advisers to the Nation on Science, Engineering, and Medicine

The nation turns to the National Academies—National Academy of Sciences, National Academy of Engineering, Institute of Medicine, and National Research Council—for independent, objective advice on issues that affect people's lives worldwide.
www.national-academies.org

ISBN 978-0-309-14286-1
90000
9 780309 142861